

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 13:46:59 ; Search time 15 Seconds

(without alignments)  
83.366 Million cell updates/sec

Title: US-09-674-616A-2

Sequence: 1 NQOVSPYTLKG 13

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 segs, 96191526 residues

Total number of hits satisfying chosen parameters: 246

Minimum DB seq length: 13

Maximum DB seq length: 13

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 78:.\*  
2: PIR:.\*  
3: PIR:.\*  
4: PIR:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	29.9	13	2 S47361	T-cell antigen rec
2	20	29.9	13	2 B56864	dipeptidyl-peptida
3	19	28.4	13	2 S32473	lymadFamide 3 - g
4	19	28.4	13	2 PL0157	Ig kappa chain V-I
5	18	26.9	13	2 B61458	Ig kappa chain V-I
6	18	26.9	13	2 A61458	Ig kappa chain V-I
7	17	25.4	13	1 UNBO	neurotensin - bovi
8	17	25.4	13	2 A44818	extracellular lipa
9	17	25.4	13	2 A53608	neurotensin - guin
10	17	25.4	13	2 A28505	neurotensin [valid
11	17	25.4	13	2 A61067	neurotensin - comm
12	17	25.4	13	2 A05174	tryptophyltin-13 -
13	17	25.4	13	2 PT0256	Ig heavy chain CRD
14	16	23.9	13	2 PC1149	equinacoxin 1A - s
15	16	23.9	13	2 A61514	glutathione transf
16	15	22.4	13	2 S15755	actin 7 - soybean
17	15	22.4	13	2 B26093	microbial collagen
18	15	22.4	13	2 A54326	glutathione transf
19	15	22.4	13	2 B28810	glutathione transf
20	15	22.4	13	2 S47358	T-cell antigen rec
21	15	22.4	13	2 PH1599	Ig H chain V-D-J r
22	14	20.9	13	2 A61361	bradykinin-like pe
23	14	20.9	13	2 PU0038	alpha-peptide/angi
24	14	20.9	13	2 S36668	hypothetical prote
25	14	20.9	13	2 S32471	lymadFamide 1 - g
26	14	20.9	13	2 S32472	lymadFamide 2 - g
27	14	20.9	13	2 D61458	Ig kappa chain V-I
28	14	20.9	13	2 E61458	Ig kappa chain V-I
29	14	20.9	13	2 S47362	T-cell antigen rec

30	14	20.9	13	2 S47380	T-cell antigen rec
31	14	20.9	13	2 B61233	concepcus protein
32	14	20.9	13	2 PH1595	Ig H chain V-D-J r
33	14	20.9	13	2 PH0786	T-cell receptor al
34	14	20.9	13	2 151905	collecting duct wa
35	13	19.4	13	2 S39413	tubulin beta chain
36	13	19.4	13	2 A32453	phloroglucinol red
37	13	19.4	13	2 S36887	ribosomal protein
38	13	19.4	13	2 PH0138	T-cell receptor be
39	13	19.4	13	2 B47415	mannose-1-phosphat
40	13	19.4	13	2 G37266	Ig heavy chain C r
41	13	19.4	13	2 D37267	Ig heavy chain C r
42	13	19.4	13	2 C47630	Ig kappa chain J r
43	13	19.4	13	2 G83968	hypothetical prote
44	13	19.4	13	2 H85575	hypothetical prote
45	13	19.4	13	2 A59491	epithelial dog all

#### ALIGNMENTS

##### RESULT 1

S47361

T-cell antigen receptor VJ junction beta chain - human

C/Species: Homo sapiens (man)

C/Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999

C/Accession: S47361

R/Lehner, P.J.

Submitted to the EMBL Data Library, August 1994

A/Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T

A/Reference number: S47355

A/Accession: S47351

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-13 <LRH>

A/Cross-references: EMBL:Z35665; NID:G527459; PIDN:CAA84754.1; PID:G527460

C/Keywords: T-cell receptor

Query Match 29.9%; Score 20; DB 2; Length 13;

Best Local Similarity 80.0%; Pred. No. 1.6e+03; Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 QVSPY 8

DB 5 QGSPY 9

##### RESULT 2

B56864 dipeptidyl-peptidase IV (EC 3.4.14.5) - bovine (fragment)

C/Species: Bos primigenius taurus (cattle)

C/Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 05-Jan-1996

C/Accession: B56864

R/Plakidou-Dymock, S.; McGivan, J.D.

Biochim. Biophys. Acta 1145, 105-112, 1993

A/Title: The oligomeric structure of renal aminopeptidase N from bovine brush-border mem

A/Reference number: A56864; PMID:93136203; PMID:8093665

A/Accession: B56864

A/Status: preliminary

A/Molecule type: protein

A/Residues: 1-13 <PLA>

A/Experimental source: renal brush-border membrane vesicles

C/Keywords: dipeptidylpeptide hydrolase

Query Match 29.9%; Score 20; DB 2; Length 13;

Best Local Similarity 37.5%; Pred. No. 1.6e+03; Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 SPYTLKG 13

DB 3 TPWKVLGS 10

```

RESULT 3
S32473
LymnadPamide 3 - great pond snail
C:Species: Lymnaea stagnalis (great pond snail)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C:Accession: S32473
R:Johnson, A.H.; Rehfeld, J.F.
Eur. J. Biochem. 213, 875-879, 1993
A:Title: LymnadPamides, a new family of neuropeptides from the pond snail, Lymnaea stagnalis
A:Reference number: S32471; MUID:9328777; PMID:8477756
A:Accession: S32473
A:Molecule type: protein
A:Residues: 1-13 <BRO>
A:Cross-references: FIDN:AA26364.1; PID:g299831
A:Experimental source: ganglia
C:Keywords: amidated carboxyl end; neuropeptide
F,1/Modified site: amidated carboxyl end (Phe) #status predicted

Query Match      28.4%; Score 19; DB 2; Length 13;
Best Local Similarity 42.9%; Pred. No. 2.4e+03;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      7 PYTLKG 13
      ||| : |
Db      1 PYDRISG 7

RESULT 4
PL0157
Ig kappa chain V-II region (anti-myelin-associated glycoprotein, PEC) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 07-Feb-1997
C:Accession: PL0157; C61458
R:Broutet, J.C.; Dellagi, K.; Gendron, M.C.; Chevalier, A.; Schmitt, C.; Mihaesco, E.
J. Exp. Med. 170, 1551-1558, 1989
A:Title: Expression of a public idioType by human monoclonal IGM directed to myelin-assoc
A:Reference number: A61458; MUID:90039128; PMID:2478651
A:Accession: PL0157
A:Molecule type: protein
A:Residues: 1-13 <BRO>
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-13 <BR2>
C:Comment: This protein is one of monoclonal IGM reactive with myelin-associated glycopr
C:Keywords: glycoprotein; heterotetramer; immunoglobulin

Query Match      28.4%; Score 19; DB 2; Length 13;
Best Local Similarity 80.0%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      6 SPYTL 10
      |||
Db      7 SPYTL 11

RESULT 5
B61458
Ig kappa chain V-I region (BLA) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 16-Aug-1996
C:Accession: B61458
R:Broutet, J.C.; Dellagi, K.; Gendron, M.C.; Chevalier, A.; Schmitt, C.; Mihaesco, E.
J. Exp. Med. 170, 1551-1558, 1989
A:Title: Expression of a public idioType by human monoclonal IGM directed to myelin-assoc
A:Reference number: A61458; MUID:90039128; PMID:2478651
A:Accession: B61458
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-13 <BRO>
C:Keywords: heterotetramer; immunoglobulin

Query Match      26.9%; Score 18; DB 2; Length 13;

```

```

      Best Local Similarity 80.0%; Pred. No. 3.8e+03;
      Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      6 SPYTL 10
      |||
Db      7 SPYTL 11

RESULT 6
A61458
Ig kappa chain V-I region (BOU) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 16-Aug-1996
C:Accession: A61458; PL0156
R:Broutet, J.C.; Dellagi, K.; Gendron, M.C.; Chevalier, A.; Schmitt, C.; Mihaesco, E.
J. Exp. Med. 170, 1551-1558, 1989
A:Title: Expression of a public idioType by human monoclonal IGM directed to myelin-assoc
A:Reference number: A61458; MUID:90039128; PMID:2478651
A:Accession: A61458
A:Molecule type: protein
A:Residues: 1-13 <BRO>
C:Comment: This protein is one of monoclonal IGM reactive with myeloma-associated glycopr
C:Keywords: heterotetramer; immunoglobulin

Query Match      26.9%; Score 18; DB 2; Length 13;
Best Local Similarity 80.0%; Pred. No. 3.8e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      6 SPYTL 10
      |||
Db      7 SPYTL 11

RESULT 7
UNB0
neurotensin - bovine
C:Species: Bos primigenius taurus (cattle)
C>Date: 18-Dec-1981 #sequence_revision 18-Dec-1981 #text_change 24-Feb-1995
C:Accession: A01420
R:Carraway, R.; Leeman, S.E.
J. Biol. Chem. 250, 1907-1911, 1975
A:Title: The amino acid sequence of a hypothalamic peptide, neurotensin.
A:Reference number: A92172; MUID:75095678; PMID:1167549
A:Accession: A01420
A:Molecule type: protein
A:Residues: 1-13 <CAR>
A:Experimental source: hypothalamus
R:Carraway, R.; Leeman, S.E.
J. Biol. Chem. 250, 1912-1918, 1975
A:Title: The synthesis of neurotensin.
A:Reference number: A92173; MUID:75095679; PMID:1112838
A:Contents: annotation; synthesis
A:Note: a tridecapeptide chemically and pharmacologically identical with the natural pep
C:Superfamily: neurotensin
C:Keywords: neuropeptide; pyroglutamic acid
F,1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match      25.4%; Score 17; DB 1; Length 13;
Best Local Similarity 75.0%; Pred. No. 5.9e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      7 PYTL 10
      |||
Db      10 PYTL 13

RESULT 8
A44818
extracellular lipase - Pseudomonas aeruginosa (fragment)
C:Species: Pseudomonas aeruginosa
C>Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 31-Oct-1997
C:Accession: A44818
R:Gilbert, E.J.; Cornish, A.; Jones, C.W.

```

J. Gen. Microbiol. 137, 2223-2229, 1991  
 A>Title: Purification and properties of extracellular lipase from *Pseudomonas aeruginosa*  
 A:Reference number: A44818; MUID:92085040; PMID:1748875  
 A:Accession: A44818  
 A>Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-13 <GIL>  
 A:Experimental source: strain TE3285  
 A>Note: sequence extracted from NCBI backbone (NCBIF:70395)  
 C:Superfamily: Pseudomonas triacylglycerol lipase

Query Match 25.4%; Score 17; DB 2; Length 13;  
 Best Local Similarity 57.1%; Pred. No. 5.9e+03;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 SPYTLK 12  
 |||  
 1 STYQTK 7

## RESULT 9

A53608  
 neurotensin - guinea pig  
 C:Species: *Cavia porcellus* (guinea pig)  
 C>Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 02-Sep-2000  
 C:Accession: A53608  
 R:Shaw, C.; Thim, L.; Conlon, J.M.  
 FEBS Lett. 202, 187-191, 1986  
 A>Title: [Ser(7)]neurotensin: isolation from guinea pig intestine.  
 A:Reference number: A53608; MUID:86248085; PMID:3087775  
 A:Accession: A53608  
 A>Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 1-13 <SHA>  
 C:Superfamily: neurotensin  
 C:Keywords: neuropeptide; pyroglutamic acid  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 25.4%; Score 17; DB 2; Length 13;  
 Best Local Similarity 75.0%; Pred. No. 5.9e+03;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 PYTL 10  
 |||  
 10 PYTL 13

## RESULT 10

A28505  
 neurotensin [validated] - chicken  
 C:Species: *Gallus gallus* (chicken)  
 C>Date: 19-Nov-1988 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
 C:Accession: A28505  
 R:Iwabuchi, H.; Komori, S.; Ohashi, H.; Kimura, S.  
 Jpn. J. Pharmacol. 44, 455-459, 1987  
 A>Title: The amino acid sequence of a smooth muscle-contracting peptide from chicken rectum  
 A:Reference number: A28505; MUID:88063566; PMID:3682409  
 A:Accession: A28505  
 A:Molecule type: protein  
 A:Residues: 1-13 <IWA>  
 A:Experimental source: rectum  
 C:Comment: The peptide isolated from rectum is identical with chicken neurotensin and has  
 C:Superfamily: neurotensin  
 C:Keywords: hormone; neuropeptide; pyroglutamic acid  
 F:1-13/Product: neurotensin #status experimental <NTS>  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 25.4%; Score 17; DB 2; Length 13;  
 Best Local Similarity 75.0%; Pred. No. 5.9e+03;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 PYTL 10  
 |||

Db 10 PYTL 13

## RESULT 11

A61067  
 neurotensin - common frog  
 C:Species: *Rana temporaria* (common frog)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 05-Aug-1994  
 C:Accession: A61067  
 R:Shaw, C.; McKay, D.M.; Halton, D.W.; Thim, L.; Buchanan, K.D.  
 Regul. Pept. 38, 23-31, 1992  
 A>Title: Isolation and primary structure of an amphibian neurotensin.  
 A:Reference number: A61067; MUID:92245104; PMID:1574601  
 A:Accession: A61067  
 A:Molecule type: protein  
 A:Residues: 1-13 <SHA>  
 A>Note: this peptide was identified in brain, intestine, and rectum, but not in stomach  
 C:Superfamily: neurotensin  
 C:Keywords: brain; intestine; neuropeptide; pyroglutamic acid  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 25.4%; Score 17; DB 2; Length 13;  
 Best Local Similarity 75.0%; Pred. No. 5.9e+03;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 PYTL 10  
 |||  
 10 PYTL 13

## RESULT 12

A05174  
 tryptophyllin-13 - Rohde's leaf frog  
 C:Species: *Phyllomedusa rohdei* (Rohde's leaf frog)  
 C>Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 18-Aug-2000  
 C:Accession: A05174  
 R:Montecucchi, P.C.; Gozzini, L.; Erspamer, V.  
 Int. J. Pept. Protein Res. 27, 175-182, 1986  
 A:Reference number: A05174  
 A:Accession: A05174  
 A:Molecule type: protein  
 A:Residues: 1-13 <MON>  
 C:Superfamily: unassigned animal peptides  
 C:Keywords: pyroglutamic acid; skin  
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 25.4%; Score 17; DB 2; Length 13;  
 Best Local Similarity 60.0%; Pred. No. 5.9e+03;  
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 QVSPY 8  
 |||  
 1 QKPY 5

## RESULT 13

PT0256  
 Ig heavy chain CDR3 region (clone 2-115C) - human (fragment)  
 C:Species: *Homo sapiens* (man)  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996  
 C:Accession: PT0256  
 R:Yanada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.  
 J. Exp. Med. 173, 395-407, 1991  
 A>Title: Preferential utilization of specific immunoglobulin heavy chain diversity and junctional diversity in the CDR3 region of the heavy chain of a human IgG1 antibody  
 A:Reference number: PT0222; MUID:91108337; PMID:1899102  
 A:Accession: PT0256  
 A:Molecule type: DNA  
 A:Residues: 1-13 <YAM>  
 A:Experimental source: B lymphocyte  
 C:Keywords: heterotetramer; immunoglobulin

Query Match 25.4%; Score 17; DB 2; Length 13;  
 Best Local Similarity 50.0%; Pred. No. 5.9e+03;

Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 YTLKLG 13  
| : | |  
Db 4 YDMLTG 9

## RESULT 14

PC1149  
equinatoxin 1A - sea anemone (Actinia equina) (fragment)  
C:Species: Actinia equina  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: PC1149  
R:Komatsu, S.; Furukawa, K.; Abe, K.; Hirano, H.; Ueda, M.  
Chem. Pharm. Bull. 40, 2873-2875, 1992  
A:Title: Isolation and characterization of equinatoxins from the sea anemone Actinia equina  
A:Reference number: PC1149; MUID:93099631; PMID:1361161  
A:Accession: PC1149  
A:Molecule type: protein  
A:Residues: 1-13 <KOM>  
C:Keywords: toxin

Query Match 23.9%; Score 16; DB 2; Length 13;  
Best Local Similarity 60.0%; Pred. No. 9.1e+03;  
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NOEQV 5  
| : |  
Db 4 NOAEV 8

## RESULT 15

A61514  
glutathione transferase (EC 2.5.1.18), 26k - fluke (Schistosoma japonicum) (fragment)  
C:Species: Schistosoma japonicum  
C:Date: 15-Oct-1994 #sequence\_revision 15-Oct-1994 #text\_change 07-May-1999  
C:Accession: A61514  
R:Mitchell, G.F.  
Mol. Biochem. Parasitol. 27, 249-256, 1988  
A:Title: Expression of an enzymatically active parasite molecule in Escherichia coli: Schistosoma japonicum  
A:Reference number: A61514; MUID:88142994; PMID:3278228  
A:Accession: A61514  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-13 <MIT>  
C:Keywords: transferase

Query Match 23.9%; Score 16; DB 2; Length 13;  
Best Local Similarity 50.0%; Pred. No. 9.1e+03;  
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 YTLKLG 13  
| : | |  
Db 6 YWKIKG 11

Search completed: October 1, 2004, 13:48:33  
Job time : 16 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 13:46:58 ; Search time 9 Seconds

(Without alignments)  
75.212 Million cell updates/sec

Title: US-09-674-616A-2

Sequence: 1 NQEQVSPYTLTKG 13

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 96

Minimum DB seq length: 13  
Maximum DB seq length: 13

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_42:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	29.9	13	1	IDHA CANFA
2	19	28.4	13	1	CHEP PARID
3	19	28.4	13	1	NP3 LYMTST
4	18	26.9	13	1	FR12 PRA
5	17	25.4	13	1	NEUT BUPMA
6	17	25.4	13	1	NEUT CAVPO
7	17	25.4	13	1	NEUT CHICK
8	17	25.4	13	1	NEUT RANTE
9	17	25.4	13	1	TY13 PHYRO
10	16	23.9	13	1	PED1 HYDAT
11	16	23.9	13	1	PROX ORYSA
12	16	23.9	13	1	SA2B ONCMY
13	16	23.9	13	1	SA2B ONCMY
14	15	22.4	13	1	ACT7 SOYBN
15	15	22.4	13	1	E121 LITRU
16	15	22.4	13	1	E122 LITRU
17	15	22.4	13	1	UN12 CLOPA
18	15	22.4	13	1	UNP1 UPEIN
19	14	20.9	13	1	BRK PARID
20	14	20.9	13	1	CRBL VESMA
21	14	20.9	13	1	MP1 MITOC
22	14	20.9	13	1	NP1 LYMTST
23	14	20.9	13	1	NP2 LYMTST
24	14	20.9	13	1	SODM CANFA
25	14	20.9	13	1	TEMG RANTE
26	14	20.9	13	1	YPE2 LACIC
27	13	19.4	13	1	FTBB RABIT
28	13	19.4	13	1	PHGR EUBOX
29	13	19.4	13	1	UN10 CLOPA
30	12	17.9	13	1	CRBL VESAN
31	12	17.9	13	1	CRBL VESCR
32	12	17.9	13	1	CRBL VESXA
33	12	17.9	13	1	PSAE_PEA

34	12	17.9	13	1	PSAE_PEA	P17229	Pisum sativ
35	12	17.9	13	1	UHA3 CANFA	P56535	canis fam1
36	12	17.9	13	1	VGL6 BACSU	P80867	bacillus su
37	11	16.4	13	1	AUI1 LITRA	P82366	litorea ran
38	11	16.4	13	1	AUI2 LITRA	P82367	litorea ran
39	11	16.4	13	1	BPE1 BOTJA	P01020	botrops ja
40	11	16.4	13	1	CRBL ICASP	P17237	icaria sp.
41	11	16.4	13	1	CRBL VESLE	P17235	vespula lew
42	11	16.4	13	1	CRTC BOVIN	P28489	bos taurus
43	11	16.4	13	1	CRTC RANES	P31832	rana esculen
44	11	16.4	13	1	EP65 HUMAN	P54963	homo sapien
45	11	16.4	13	1	GER1 HORVU	P28525	hordeum vul

## ALIGNMENTS

## RESULT 1

IDHA CANFA STANDARD; PRT; 13 AA.

AC P54836;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-0UL-1998 (Rel. 36, Last annotation update)  
DE Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial  
DE (EC 1.1.1.41) (Isocitric dehydrogenase) (NAD+-specific IDH)  
DE (Fragment).  
GN IDH3A.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE.

RC TISSUE=Heart;  
RX MEDLINE=98163340; PubMed=9504812;  
RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
dog heart proteins."  
RL Electrophoresis 18:2795-2802(1997).  
CC -1- CATALYTIC ACTIVITY: Isocitrate + NAD(+) = 2-oxoglutarate + CO(2) +  
NADH.  
CC -1- SUBUNIT: Heterooligomer of subunits alpha, beta, and gamma in the  
CC -1- APPARENT RATIO OF 2:1:1 (By similarity).  
CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
CC -1- SIMILARITY: Belongs to the isocitrate and isopropylmalate  
CC dehydrogenases family.

DR HSC-2DPAGE; P54836; DOG.  
DR InterPro; IPR001804; Isoch.  
DR PROSITE; PS00470; IDH IMDH; PARTIAL.  
KW Oxidoreductase; NAD; Tricarboxylic acid cycle; Mitochondrion.  
FT NON TER 13  
SQ SEQUENCE 13 AA; 1356 MW; 9ABFBC2B2A34B2D1 CRC64;  
Query Match 29.9%; Score 20; DB 1; Length 13;  
Best Local Similarity 40.0%; Pred. No. 8.3e+02;  
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 QVSPYTLTKG 13

DB 3 EVQVTTLTPG 12

## RESULT 2

CHEP PARID STANDARD; PRT; 13 AA.

AC P42718;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Chemotactic peptide.  
OS Parapolybia indica.  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Vespoidea;  
 OC Vespiidae; Polistinae; Parapolybia.  
 OX NCBI\_TaxID=31921;  
 RN [1]  
 RP SEQUENCE.  
 RP TISSUE=Venom;  
 RA Toki T., Yasuhara T., Nakajima T.;  
 RT "Isolation and sequential analysis of peptides on the venom sac of  
 Parapolybia indica.";  
 RL Bisel Dobutau 39:105-111(1988).  
 KW Chemotaxis; Amidation.  
 FT MOD RES 13  
 SQ SEQUENCE 13 AA; 1298 MW; 5C950CE839D5873 CRC64;  
 QY Query Match 28.4%; Score 19; DB 1; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 10 LKLG 13  
 4 LKLG 7

RESULT 3  
 NP3\_LYMST STANDARD; PRT; 13 AA.  
 AC P80180;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-JUL-1993 (Rel. 26, Last annotation update)  
 DE Lyman-DF-amide 3.  
 OS Lymanaea stagnalis (Great pond snail).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
 OC Lymanaeoidea; Lymanaeidae; Lymanaea.  
 OX NCBI\_TaxID=6523;  
 RN [1]  
 RP SEQUENCE.  
 RP TISSUE=Ganglion;  
 RX MEDLINE=9328777; PubMed=8477756;  
 RA Johnson A.H., Rehfeld J.F.;  
 RT "LymanadFamides, a new family of neuropeptides from the pond snail,  
 Lymanaea stagnalis. Clue to cholecystokinin immunoreactivity in  
 invertebrates?";  
 RL But. J. Biochem. 213:875-879(1993).  
 CC -1- SIMILARITY: RELATED TO THE CHOLECYSTOKININ (CCK) FAMILY.  
 DR PIR; S32473; S32473.  
 KW Neuropeptide; Amidation.  
 FT MOD RES 13  
 FT UNSURE 12  
 SQ SEQUENCE 13 AA; 1462 MW; 9CA07BA3F5D5B865 CRC64;  
 QY Query Match 28.4%; Score 19; DB 1; Length 13;  
 Best Local Similarity 42.9%; Pred. No. 1.3e+03;  
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 Db 7 PYTLKG 13  
 1 PYDRISG 7

RESULT 4  
 FR12\_PEA STANDARD; PRT; 13 AA.  
 ID FR12\_PEA  
 AC P83445;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ferritin 2, chloroplast (Fragment).  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
 OX NCBI\_TaxID=3888;

RN [1]  
 RP SEQUENCE, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.  
 RP STRAIN=cv. Laxton's Progress; TISSUE=leaf;  
 RA Shingler R., McCarty R.E.;  
 RL Submitted (SEP-2002) to Swiss-Prot.  
 CC -1- FUNCTION: Ferritin is an intracellular molecule that stores iron  
 in a soluble, nontoxic, readily available form. The functional and  
 molecule, which is composed of 24 chains, is roughly spherical and  
 contains a central cavity into which the polymeric ferric iron  
 core is deposited.  
 CC -1- SUBCELLULAR LOCATION: Chloroplast (inner envelope membrane), and  
 other plastids.  
 CC -1- TISSUE SPECIFICITY: Leaves.  
 CC -1- MISCELLANEOUS: On the 2D-gel the determined pI of this protein is:  
 4.74, its MW is: 25.7 kDa.  
 CC -1- SIMILARITY: Belongs to the ferritin family.  
 DR InterPro: IPR001519; Ferritin.  
 DR InterPro: IPR009040; Ferritin-like.  
 DR PROSITE: PS00204; FERRITIN\_2; PARTIAL.  
 DR PROSITE: PS00540; FERRITIN\_1; PARTIAL.  
 DR PROSITE: PS50905; FERRITIN-LIKE; PARTIAL.  
 KW Iron storage; Iron; Metal-binding; Chloroplast.  
 FT NON\_TER 13  
 SQ SEQUENCE 13 AA; 1246 MW; 26C9DC25F34ADC7 CRC64;  
 QY Query Match 26.3%; Score 18; DB 1; Length 13;  
 Best Local Similarity 80.0%; Pred. No. 2e+03;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 6 SPYTL 10  
 5 SPATL 9

RESULT 5  
 NEUT\_BUPMA STANDARD; PRT; 13 AA.  
 ID NEUT\_BUPMA  
 AC P81796;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Neurotensin (NT).  
 OS Bufo marinus (Giant toad) (Cane toad).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Bufonidae; Bufo.  
 OX NCBI\_TaxID=8386;  
 RN [1]  
 RP SEQUENCE, AND SYNTHESIS.  
 RP TISSUE=Small intestine;  
 RX MEDLINE=99000115; PubMed=9786176;  
 RA Warner F.J., Butcher E., Carriaway R., Conlon J.M.;  
 RT "Purification, characterization, and spasmogenic activity of  
 neurotensin from the toad Bufo marinus.";  
 RL Peptides 19:1255-1261(1998).  
 CC -1- FUNCTION: Smooth muscle-contracting peptide. Peptide action is not  
 affected by tetrodotoxin, but is slightly mediated through the  
 release of acetylcholine.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the neurotensin family.  
 KW Vasoactive; Pyrrolidone carboxylic acid.  
 FT MOD RES 1  
 SQ SEQUENCE 13 AA; 1515 MW; 56BD9F3A5410DD3 CRC64;  
 QY Query Match 25.4%; Score 17; DB 1; Length 13;  
 Best Local Similarity 75.0%; Pred. No. 3.2e+03;  
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 7 PYTL 10  
 10 PYTL 13

RESULT 6

```

NEUT_CAVPO
ID NEUT_CAVPO STANDARD; PRT; 13 AA.
AC P32560;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurotensin (NT).
GN NTS.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE.
RC TISSUE=Small intestine;
RX MEDLINE=86248085; PubMed=3087775;
RA Shaw C., Thim L., Conlon J.M.;
RT "[Ser?]neurotensin: isolation from guinea pig intestine.";
RL FEBS Lett. 202:187-192 (1986).
CC -1- FUNCTION: Smooth muscle-contracting peptide.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the neurotensin family.
DR PIR; A53608; A53608.
KW Vasactive; Pyrrolidone carboxylic acid.
FT MOD RES 1
SQ SEQUENCE 13 AA; 1680 MW; 4C831464C4115B3 CRC64;

Query Match
Best Local Similarity 25.4%; Score 17; DB 1; Length 13;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 PYTL 10
DB 10 PYTL 13

RESULT 7
NEUT_CHICK
ID NEUT_CHICK STANDARD; PRT; 13 AA.
AC P13724;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurotensin (NT).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=88063566; PubMed=3682409;
RA Iwabuchi H., Komori S., Ohashi H., Kimura S.;
RT "The amino acid sequence of a smooth muscle-contracting peptide from
RT chicken rectum. Identity to chicken neurotensin.";
RL Jpn. J. Pharmacol. 44:455-459 (1987).
CC -1- FUNCTION: Smooth muscle-contracting peptide.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the neurotensin family.
DR PIR; A28505; A28505.
KW Vasactive; Pyrrolidone carboxylic acid.
FT MOD RES 1
SQ SEQUENCE 13 AA; 1608 MW; 4C949E714C410D3 CRC64;

Query Match
Best Local Similarity 25.4%; Score 17; DB 1; Length 13;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 PYTL 10
DB 10 PYTL 13

```

```

RESULT 8
NEUT_RANTE
ID NEUT_RANTE STANDARD; PRT; 13 AA.
AC P41536;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurotensin (NT).
OS Rana temporaria (European common frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8407;
RN [1]
RP SEQUENCE.
RX MEDLINE=92245104; PubMed=1574601;
RA Shaw C., McKay D.M., Halton D.W., Thim L., Buchanan K.D.;
RT "Isolation and primary structure of an amphibian neurotensin.";
RL Regul. Pept. 38:23-31 (1992).
CC -1- FUNCTION: Smooth muscle-contracting peptide.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Identified in brain, intestine, and rectum,
CC but not in stomach or skin.
CC -1- SIMILARITY: Belongs to the neurotensin family.
DR PIR; A61067; A61067.
KW Vasactive; Pyrrolidone carboxylic acid.
FT MOD RES 1
SQ SEQUENCE 13 AA; 1569 MW; 56A53D69EF410D3 CRC64;

Query Match
Best Local Similarity 25.4%; Score 17; DB 1; Length 13;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 PYTL 10
DB 10 PYTL 13

RESULT 9
TY13_PHYRO
ID TY13_PHYRO STANDARD; PRT; 13 AA.
AC P04056;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tryptophyllin-13.
OS Phyllomedusa rohdei (Rohde's leaf frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;
OC Phyllomedusinae; Phyllomedusa.
OX NCBI_TaxID=8394;
RN [1]
RP SEQUENCE.
RA Montecucchi P.C., Gozzini L., Erspamer V.;
RT "Primary structure determination of a tryptophan-containing
RT tridecapeptide from Phyllomedusa rohdei.";
RL Int. J. Pept. Protein Res. 27:175-182 (1986).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Skin.
DR PIR; A05174; A05174.
KW Amphibian defense peptide; Pyrrolidone carboxylic acid.
FT MOD RES 1
SQ SEQUENCE 13 AA; 1646 MW; 33BF33A212227773 CRC64;

Query Match
Best Local Similarity 25.4%; Score 17; DB 1; Length 13;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 QVSPY 8
DB 1 QEKPY 5

RESULT 10

```

```

PEDI_HYDAT
ID PEDI_HYDAT STANDARD; PRT; 13 AA.
AC P80578;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE Pedin. 1996 (Rel. 34, Last annotation update)
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydroids; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxId=6087;
RN [1]
RP SEQUENCE.
RX MEDLINE=66232307; PubMed=6674432;
RA Hoffmeister S.A.H.;
RT "Isolation and characterization of two new morphogenetically active
RL peptides from Hydra vulgaris."
RL Development 122:1941-1948(1996).
CC -1- FUNCTION: Morphogenetically active peptide. Active in foot
development.
KW Morphogen. 13 AA; 1512 MW; 6F5266E3F37632CB CRC64;
SQ SEQUENCE 13 AA; 1362 MW; 0A3022E0E52C66B CRC64;

Query Match 23.9%; Score 16; DB 1; Length 13;
Best Local Similarity 37.5%; Pred. No. 5e+03;
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 EQVSPYTL 10
Db 1 EELRPEVL 8

RESULT 11
PROX ORYSA STANDARD; PRT; 13 AA.
ID PROX ORYSA STANDARD; PRT; 13 AA.
AC P83647;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable profilin LP04 (Fragments).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxId=4530;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. Indica-IR64; TISSUE=Panicle;
RA Hosseini Salekdeh S.G.; Bennett J.;
RT "Proteome analysis of rice panicle."
RL Submitted (JUL-2003) to Swiss-Prot.
CC -1- FUNCTION: Binds to actin and affects the structure of the
cytoskeleton. At high concentrations, profilin prevents the
polymerization of actin, whereas it enhances it at low
concentrations. By binding to PIP2, it inhibits the formation of
F-actin and DG (by similarity).
CC -1- SUBUNIT: Occurs in many kinds of cells as a complex with monomeric
actin in a 1:1 ratio.
CC -1- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
protein is: 4.4, its MW is: 14.0 kDa.
CC -1- SIMILARITY: Belongs to the profilin family.
DR PROSITE; PS00414; PROFILIN; PARTIAL.
KW Actin-binding; Cytoskeleton; Multigene family.
FT NON TER 1
FT NON TER 1
FT NON TER 5
FT NON TER 6
SQ SEQUENCE 13 AA; 1362 MW; 0A3022E0E52C66B CRC64;

Query Match 23.9%; Score 16; DB 1; Length 13;
Best Local Similarity 33.3%; Pred. No. 5e+03;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 YTLKLG 13

```

```

Db 6 YWVIQG 11

RESULT 12
SA2B ONCMY STANDARD; PRT; 13 AA.
ID SA2B ONCMY STANDARD; PRT; 13 AA.
AC P82238;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Salmocidin 2A (Fragment).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxId=8022;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Sera;
RA Henry M.A.; Secombes C.J.;
RT "Purification and partial characterization of antibacterial peptides
RL from rainbow trout, Oncorhynchus mykiss."
RL Submitted (DEC-1999) to Swiss-Prot.
CC -1- FUNCTION: Antibacterial activity against Gram-negative bacteria.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma serum.
KW Antibiotic.
FT NON TER 13
FT NON TER 13
FT NON TER 13
SQ SEQUENCE 13 AA; 1416 MW; 5FFB792AF6C45873 CRC64;

Query Match 23.9%; Score 16; DB 1; Length 13;
Best Local Similarity 75.0%; Pred. No. 5e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 LKLG 13
Db 4 VLKG 7

RESULT 13
SA2B ONCMY STANDARD; PRT; 13 AA.
ID SA2B ONCMY STANDARD; PRT; 13 AA.
AC P82239;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Salmocidin 2B (Fragment).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxId=8022;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Sera;
RA Henry M.A.; Secombes C.J.;
RT "Purification and partial characterization of antibacterial peptides
RL from rainbow trout, Oncorhynchus mykiss."
RL Submitted (DEC-1999) to Swiss-Prot.
CC -1- FUNCTION: Antibacterial activity against Gram-negative bacteria.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Plasma serum.
KW Antibiotic.
FT NON TER 13
FT NON TER 13
FT NON TER 13
SQ SEQUENCE 13 AA; 1400 MW; 5FFB792AF6C45873 CRC64;

Query Match 23.9%; Score 16; DB 1; Length 13;
Best Local Similarity 75.0%; Pred. No. 5e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 LKLG 13

```



```

Db          4 VLKG 7

RESULT 14
ACT7_SOYBN  STANDARD;      PRT;      13 AA.
ID ACT7_SOYBN
AC P15987;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Actin 7 (Fragment).
GN SACT.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Wayne;
RX MEDLINE=91346640; PubMed=2102831;
RA Pearson L., Meagher R.B.;
RT "Diverse soybean actin transcripts contain a large intron in the 5'
RT untranslated leader: structural similarity to vertebrate muscle actin
RT genes."
RL Plant Mol. Biol. 14:513-526(1990).
CC -!- FUNCTION: Actins are highly conserved proteins that are involved
CC in various types of cell motility and are ubiquitously expressed
CC in all eukaryotic cells.
CC -!- FUNCTION: Essential component of cell cytoskeleton; plays an
CC important role in cytoplasmic streaming, cell shape determination,
CC cell division, organelle movement and extension growth.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- MISCELLANEOUS: There are at least 16 actin genes in soybean.
CC -!- SIMILARITY: Belongs to the actin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X17120; CAA34980.1; -.
DR PIR, S15755; S15755.
DR InterPro, IPR004001; Actin.
DR InterPro, IPR004000; Actin like.
DR PROSITE, PS00406; ACTINS_1; PARTIAL.
DR PROSITE, PS00432; ACTINS_2; PARTIAL.
DR PROSITE, PS01132; ACTINS_ACT LIKE; PARTIAL.
KW Structural protein; Multigene family.
FT NON TER 13
SQ SEQUENCE 13 AA; 1420 MM; 8BEFF3C36D4FD05A CRC64;

Query Match 22.4%; Score 15; DB 1; Length 13;
Best Local Similarity 40.0%; Pred. No. 7.8e+03;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 EQVSP 7
| : |
| : |
Db 5 EDIOP 9

RESULT 15
E121_LITRU  STANDARD;      PRT;      13 AA.
ID E121_LITRU
AC P82097;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Electrin 2.1.
OS Litorea rubella (Desert tree frog).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hylidae; Hylidae;
OC Pelodyadinae; Litorea.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Mabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian burzing tree frog
RT Litorea electrica. Comparison with the skin peptides from Litorea
RT rubella."
RL Aust. J. Chem. 52:639-645(1999).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Amidation.
FT MOD RES 13
SQ SEQUENCE 13 AA; 1599 MM; C1808FF326F57322 CRC64;

Query Match 22.4%; Score 15; DB 1; Length 13;
Best Local Similarity 50.0%; Pred. No. 7.8e+03;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOEQ 4
| : |
| : |
Db 1 NEEZ 4

Search completed: October 1, 2004, 13:48:13
Job time : 11 secs

```

Blank Sheet

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 13:46:59 ; Search time 37 Seconds

(without alignments)  
110,858 Million cell updates/sec

Title: US-09-674-616A-2  
Perfect score: 67  
Sequence: 1 NQEQVSPYTLKG 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 436

Minimum DB seq length: 13  
Maximum DB seq length: 13

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	32.8	13	4 Q9UEE2	Q9uee2 homo sapien
2	19	28.4	13	2 Q51605	Q51605 escherichia
3	19	28.4	13	2 Q9RG00	Q9rg00 mycoplasma
4	19	28.4	13	4 Q14890	Q14890 homo sapien
5	18	26.9	13	2 Q48357	Q48357 escherichia
6	17	26.4	13	2 Q50476	Q50476 mycobacteri
7	16	23.9	13	3 Q93824	Q93824 candida tro
8	16	23.9	13	10 Q9FSA8	Q9fsa8 silene bacc
9	16	23.9	13	10 Q9FS8N1	Q9fsn1 hordeum vul
10	16	23.9	13	10 Q9FS94	Q9fs94 silene pent
11	15	22.4	13	2 Q47693	Q47693 escherichia
12	15	22.4	13	2 Q9RFZ4	Q9rfz4 mycoplasma
13	15	22.4	13	11 Q9QW45	Q9qw45 rattus sp.
14	15	22.4	13	11 Q9QY46	Q9qy46 mus sp. zfp
15	15	22.4	13	11 Q88176	Q88176 mus musculu
16	14	20.9	13	2 Q9R8R9	Q9r8r9 streptococc

17	14	20.9	13	2 Q31295	Q31295 buchnera ap
18	14	20.9	13	4 Q96Q66	Q96q66 homo sapien
19	14	20.9	13	4 Q86V06	Q86v06 homo sapien
20	14	20.9	13	5 Q61340	Q61340 panulirus i
21	14	20.9	13	8 Q8WEJ9	Q8wej9 ginkgo bilo
22	14	20.9	13	10 P82432	P82432 nicotiana t
23	14	20.9	13	11 Q8CJ33	Q8cj33 mus musculu
24	14	20.9	13	11 Q80Y03	Q80y03 rattus sp.
25	14	20.9	13	13 P82848	P82848 rana pipien
26	13	19.4	13	3 P87031	P87031 saccharomyc
27	13	19.4	13	4 Q9UEE3	Q9uee3 homo sapien
28	13	19.4	13	4 Q14461	Q14461 homo sapien
29	13	19.4	13	5 P83567	P83567 sepi offic
30	13	19.4	13	8 Q8WGC2	Q8wgc2 isochloes p
31	13	19.4	13	8 Q85LH3	Q85lh3 plectura m
32	13	19.4	13	12 Q9E1V2	Q9e1v2 hepatitis b
33	13	19.4	13	12 Q9E1U6	Q9e1u6 hepatitis b
34	13	19.4	13	12 Q9E1U5	Q9e1u5 hepatitis b
35	13	19.4	13	12 Q9E1U0	Q9e1u0 hepatitis b
36	13	19.4	13	12 Q9E1T9	Q9e1t9 hepatitis b
37	13	19.4	13	12 Q9E1T8	Q9e1t8 hepatitis b
38	13	19.4	13	12 Q9E1U4	Q9e1u4 hepatitis b
39	13	19.4	13	12 Q9E1U9	Q9e1u9 hepatitis b
40	13	19.4	13	12 Q9WMG5	Q9wmg5 sigma virus
41	13	19.4	13	12 Q9E1U4	Q9e1u4 hepatitis b
42	13	19.4	13	12 Q9E1V4	Q9e1v4 hepatitis b
43	13	19.4	13	12 Q9E1U2	Q9e1u2 hepatitis b
44	13	19.4	13	12 Q86129	Q86129 vesicular s
45	13	19.4	13	12 Q86129	Q86129 vesicular s

## ALIGNMENTS

## RESULT 1

Q9UEE2 PRELIMINARY; PRT; 13 AA.  
AC Q9UEE2;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE ERGB transcription factor (fragment).  
GN FLIL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Obata K., Hiraga H., Nojima T., Yoshida M.C., Abe S.;  
RT "Molecular analysis on the breakpoint region of a t(11:22)  
RT translocation in Ewing's sarcoma";  
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AB012625; BAA32806.1; .  
FT NON\_TER  
FT NON\_TER  
SQ SEQUENCE 13 AA; 1541 MW; 2C677798CB566AB7 CRC64;

Query Match 32.8%; Score 22; DB 4; Length 13;  
Best Local Similarity 50.0%; Pred. No. 3e+03; 3; Indels 0; Gaps 0;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 QVSPYTL 11  
DB 5 QPDYQIL 12

RESULT 2  
ID Q51605 PRELIMINARY; PRT; 13 AA.  
AC Q51605;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

```

DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)
DE CE1E7 protein (Fragment).
GN CE1E7.
OS Escherichia coli.
OC Plasmid ColE7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=92293113; PubMed=1603061;
RA Soong B.W., Lu F.M., Chak K.F.;
RT "Characterization of the cea gene of the ColE7 plasmid.";
RL Mol. Gen. Genet. 233:177-183(1992).
DR EMBL; M62409; AAA8085.1; -.
GO; GO:0046821; C:extrachromosomal DNA, IEA.
DR InterPro; IPR000290; Colicin_pyocin.
DR Pfam; PF01320; Colicin_Pyocin; 1.
KM plasmid.
FT NON_TER
SQ SEQUENCE 13 AA; 1501 MW; 9E75F892148CB045 CRC64;

Query Match 28.4%; Score 19; DB 2; Length 13;
Best Local Similarity 37.5%; Pred. No. 1.1e+04;
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 QEOVSPPT 9
Db 4 KNSISDPT 11

RESULT 3
Q9RG00 PRELIMINARY; PRT; 13 AA.
ID Q9RG00
AC Q9RG00;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
DE Hypothetical protein (Fragment).
OS Mycoplasma capricolum subsp. capricolum.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=40479;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=8035;
RA MEDLINE=20193983; PubMed=10727835;
RA Thiaucourt F., Lorenzon S., David A., Breard A.;
RT "Phylogeny of the Mycoplasma mycoides cluster as shown by sequencing
RT of a putative membrane protein gene.";
RL Vet. Microbiol. 72:251-268(2000).
DR EMBL; AF162995; AAF15247.1; -.
KM Hypothetical protein.
FT NON_TER
SQ SEQUENCE 13 AA; 1459 MW; 0B63638AED35573B CRC64;

Query Match 28.4%; Score 19; DB 2; Length 13;
Best Local Similarity 42.9%; Pred. No. 1.1e+04;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 SPYTLK 12
Db 4 TPYLKK 10

RESULT 4
Q14890 PRELIMINARY; PRT; 13 AA.
ID Q14890
AC Q14890;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Mucin (Fragment).
GN MUC5AC.
RN [1]

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX TISSUE=Tracheobronchial mucosa;
RA Guyonnet-Duperrat V., Audie J., Debailleul V., Laine A., Bulsine M.,
RA Zoultina-Galliege S., Pigny P., Degand P., Aubert J., Porchet N.;
RT "Characterization of the human mucin gene MUC5AC: a consensus
RT cysteine-rich domain for 11p15 mucin genes.";
RL Biochem. J. 0:0-0(1994).
DR EMBL; Z34280; CAA84034.1; -.
FT NON_TER
SQ SEQUENCE 13 AA; 1580 MW; 535DF5A5183B7767 CRC64;

Query Match 28.4%; Score 19; DB 4; Length 13;
Best Local Similarity 42.9%; Pred. No. 1.1e+04;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NQROVSP 7
Db 2 NKWQLPP 8

RESULT 5
Q48357 PRELIMINARY; PRT; 13 AA.
ID Q48357
AC Q48357;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Transcription termination factor rho (Fragment).
GN RHO.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95129907; PubMed=7828920;
RA Opperman T., Martinez A., Richardson J.P.;
RT "The ts15 mutation of Escherichia coli alters the sequence of the C-
RT terminal nine residues of Rho protein.";
RL Gene 152:133-134(1995).
DR EMBL; L34404; AAA68985.1; -.
FT NON_TER
SQ SEQUENCE 13 AA; 1531 MW; 93F7EB96FDBECB1A CRC64;

Query Match 26.9%; Score 18; DB 2; Length 13;
Best Local Similarity 42.9%; Pred. No. 1.1e+04;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NQEQVSP 7
Db 5 NDEVWTP 11

RESULT 6
Q50476 PRELIMINARY; PRT; 13 AA.
ID Q50476
AC Q50476;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE Catalase.
GN KATG.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]

```

RP SEQUENCE FROM N.A.  
 RC STRAIN-110373;  
 RA Cockerill F.R., Uh1 J.R., Temessen Z., Zhang Y., Stockman L.,  
 RA Roberts G.D., Williams D.L., Kline B.C.;  
 RT "Rapid identification of a point mutation of the *Mycobacterium*  
 RT tuberculosis catalase-peroxidase (katG) gene associated with isoniazid  
 RT resistance".  
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U06263; AAB59976.1; -.  
 SQ SEQUENCE 13 AA; 1564 MW; 2F39A45EFE994777 CRC64;

Query Match 25.4%; Score 17; DB 2; Length 13;  
 Best Local Similarity 57.1%; Pred. No. 2.7e+04;  
 Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 EQVSPYT 9  
 DB 3 EQHPPT 9

RESULT 7  
 O93824 PRELIMINARY; PRT; 13 AA.  
 AC O93824;  
 DT 01-MAY-1999 (TEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TEMBLrel. 10, last sequence update)  
 DE UDP-galactose-4-epimerase (Fragment).  
 GN GAL10.  
 OS Candida tropicalis (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5482;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PK233;  
 RA Kanai T., Ueda M., Tanaka A.;  
 RT "Candida tropicalis promoter region of GAL1 and GAL10 gene."  
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB019434; BAA34351.1; -.  
 FT NON\_TER 1 13  
 FT NON\_TER 13 13  
 SQ SEQUENCE 13 AA; 1312 MW; 0B559D6AE18BC720 CRC64;

Query Match 23.9%; Score 16; DB 3; Length 13;  
 Best Local Similarity 50.0%; Pred. No. 4.2e+04;  
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 YTLKG 13  
 DB 5 YILVTG 10

RESULT 8  
 O9FS8 PRELIMINARY; PRT; 13 AA.  
 ID O9FS8  
 AC O9FS8;  
 DT 01-MAR-2001 (TEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TEMBLrel. 16, last sequence update)  
 DT 01-MAR-2001 (TEMBLrel. 16, last annotation update)  
 DE RNA polymerase II (Fragment).  
 GN RPB2.  
 OS Silene bacillifera (Berry catchfly) (Cucubalus bacillifer).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Caryophyllaceae; Silene.  
 OX NCBI\_TaxID=54818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Popp M., Oxelman B.;  
 RT "Inferring the history of the polyploid Silene aegaea  
 RT (Caryophyllaceae) using nuclear and chloroplast DNA sequence data."  
 RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ296139; CAC13015.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 13 13  
 SQ SEQUENCE 13 AA; 1437 MW; D931E2CF1AB6D1B2 CRC64;

Query Match 23.9%; Score 16; DB 10; Length 13;  
 Best Local Similarity 50.0%; Pred. No. 4.2e+04;  
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 SPYT 9  
 DB 1 TPPT 4

RESULT 9  
 O9S8N1 PRELIMINARY; PRT; 13 AA.  
 ID O9S8N1  
 AC O9S8N1;  
 DT 01-MAY-2000 (TEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TEMBLrel. 13, last sequence update)  
 DT 01-DEC-2001 (TEMBLrel. 19, last annotation update)  
 DE Calmodulin (Fragment).  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Hordeum.  
 OX NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE  
 RX MEDLINE=94266769; PubMed=8206888;  
 RA Toda H., Yazawa M., Sakiyama F., Yeji K.;  
 RT "Amino acid sequence of calmodulin from wheat germ."  
 RT J. Biochem. 115:367-367(1994).  
 SQ SEQUENCE 13 AA; 1508 MW; 9763F31CD443DB1A CRC64;

Query Match 23.9%; Score 16; DB 10; Length 13;  
 Best Local Similarity 33.3%; Pred. No. 4.2e+04;  
 Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 EQVSPY 8  
 DB 7 EQTAEP 12

RESULT 10  
 O9FS94 PRELIMINARY; PRT; 13 AA.  
 ID O9FS94  
 AC O9FS94;  
 DT 01-MAR-2001 (TEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TEMBLrel. 16, last sequence update)  
 DT 01-MAR-2001 (TEMBLrel. 16, last annotation update)  
 DE RNA polymerase II (Fragment).  
 GN RPB2.  
 OS Silene pentelica.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Caryophyllaceae; Silene.  
 OX NCBI\_TaxID=49735;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Popp M., Oxelman B.;  
 RT "Inferring the history of the polyploid Silene aegaea  
 RT (Caryophyllaceae) using nuclear and chloroplast DNA sequence data."  
 RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ296132; CAC13024.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 13 13  
 SQ SEQUENCE 13 AA; 1437 MW; D931E2CF1AB6D1B2 CRC64;

Query Match 23.9%; Score 16; DB 10; Length 13;  
 Best Local Similarity 50.0%; Pred. No. 4.2e+04;  
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 SPYT 9  
: :  
DB 1 TRPT 4

RESULT 11  
ID Q47693 PRELIMINARY; PRT; 13 AA.  
AC Q47693:

DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE Elongation factor TU (EF-TU) (Fragment).  
GN TUBB.  
OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
NCBI\_TaxID=562;

RP SEQUENCE FROM N.A.  
RX MEDLINE=82080657; PubMed=7312036;  
RA Hudson L., Rossi J., Landy A.;  
RT "Dual function transcripts specifying tRNA and mRNA.";  
RL Nature 294:422-427(1981).

DR EMBL; X04181; CAA27777.1; -  
DR GO; GO:0003746; P:translation elongation factor activity; IEA.  
DR GO; GO:0006412; P:protein biosynthesis; IEA.  
KW Elongation factor; Protein biosynthesis.

FT NON TER 13  
SQ SEQUENCE 13 AA; 1617 MW; C433BE82A18D0B19 CRC64;

Query Match 22.4%; Score 15; DB 2; Length 13;  
Best Local Similarity 33.3%; Pred. No. 6.5e+04;  
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 EQVSPY 8  
: :  
DB 7 ERTKPH 12

RESULT 12  
ID Q9RFZ4 PRELIMINARY; PRT; 13 AA.  
AC Q9RFZ4:

DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Hypothetical protein (Fragment).  
OS Mycoplasma mycoides subsp. capri.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
NCBI\_TaxID=40477;

RP SEQUENCE FROM N.A.  
RX STRAIN=EG3;  
RX MEDLINE=20193983; PubMed=10727835;

RA Thaucourt F., Lorenzon S., David A., Breard A.;  
RT "Phylogeny of the Mycoplasma mycoides cluster as shown by sequencing  
RT of a putative membrane protein gene.";  
RL Vet. Microbiol. 72:251-268(2000).

DR EMBL; AF162998; AAF15253.1; -  
KW Hypothetical protein.  
FT NON TER 1  
SQ SEQUENCE 13 AA; 1505 MW; 0B79431F5635573B CRC64;

Query Match 22.4%; Score 15; DB 2; Length 13;  
Best Local Similarity 50.0%; Pred. No. 6.5e+04;  
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 PYTL 10  
: :  
DB 5 PYLI 8

RESULT 13  
ID Q9QW45 PRELIMINARY; PRT; 13 AA.  
AC Q9QW45:

DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE GTP-binding protein RAB16 (Fragment).

OS Rattus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;

RP SEQUENCE FROM N.A.  
RX MEDLINE=93054572; PubMed=1429617;  
RA Eiferink L.A., Anzai K., Scheller R.H.;  
RT "Rab15, a novel low molecular weight GTP-binding protein specifically  
RT expressed in rat brain.";  
RL J. Biol. Chem. 267:22693-22693(1992).

FT NON TER 13  
SQ SEQUENCE 13 AA; 1584 MW; 40BD5FE7236041A3 CRC64;

Query Match 22.4%; Score 15; DB 11; Length 13;  
Best Local Similarity 33.3%; Pred. No. 6.5e+04;  
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 QVSPYT 9  
: :  
DB 4 QIKTYS 9

RESULT 14  
ID Q9QY46 PRELIMINARY; PRT; 13 AA.  
AC Q9QY46:

DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE Zfp127 protein (Fragment).  
OS Mus sp.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10095;

RP SEQUENCE FROM N.A.  
RX MEDLINE=20056261; PubMed=10588722;  
RA Greally J.M., Gray T.A., Gabriel J.M., Song L., Zemel S.,  
RA Nicholls R.D.;

RT "Conserved characteristics of heterochromatin-forming DNA at the  
RT 15q11-q13 imprinting center.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:14430-14435(1999).

DR EMBL; AF130348; AAD56756.1; -  
FT NON TER 13  
SQ SEQUENCE 13 AA; 1356 MW; D89F5BA9529476D CRC64;

Query Match 22.4%; Score 15; DB 11; Length 13;  
Best Local Similarity 33.3%; Pred. No. 6.5e+04;  
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 QEOVSP 7  
: :  
DB 2 EESTAP 7

RESULT 15  
ID 088176 PRELIMINARY; PRT; 13 AA.  
AC 088176:

DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Neural cell adhesion molecule (Fragment).

GN NCAM1 OR NCAM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Balb-c; TISSUE=Liver;  
 RX MEDLINE=98250618; PubMed=9582442;  
 RA Kawahigashi H., Harada Y., Asano A., Nakamura M.;  
 RT "A cis-acting regulatory element that affects the alternative splicing  
 of a muscle-specific exon in the mouse NCAM gene.";  
 RL Biochim. Biophys. Acta 1397:305-315(1998).  
 DR EMBL; AB001873; BAA31275.1; -.  
 DR MGD; MGI:97281; Ncam1.  
 FT NON\_TER 1  
 FT NON\_TER 13  
 SQ SEQUENCE 13 AA; 1496 MW; CC609BF54C72D732 CRC64;

Query Match 22.4%; Score 15; DB 11; Length 13;  
 Best Local Similarity 42.9%; Pred. No. 6.5e+04;  
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 VSEPTLL 11  
 : |||  
 Db 7 LQPTCT 13

Search completed: October 1, 2004, 13:50:08  
 Job time : 39 secs

Blank Sheet



Fri Oct 1 13:54:09 2004

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 13:48:03 ; Search time 17 Seconds

(without alignments)  
39,479 Million cell updates/sec

Title: US-09-674-616A-2

Sequence: 1 NOEQVSPYTLKG 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 segs, 51625971 residues

Total number of hits satisfying chosen parameters: 6327

Minimum DB seq length: 13  
Maximum DB seq length: 13

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCUTS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	53.7	13	4	US-09-675-922-1
2	24	35.8	13	1	US-08-174-467-6
3	24	35.8	13	3	US-08-452-071-6
4	23	34.3	13	1	US-08-620-213-3
5	23	34.3	13	2	US-08-542-686-1
6	23	34.3	13	2	US-08-332-562A-69
7	22	32.8	13	1	US-08-068-947-14
8	22	32.8	13	1	US-08-068-947-15
9	22	32.8	13	3	US-08-068-947-22
10	22	32.8	13	4	US-09-468-578-11
11	22	32.8	13	4	US-09-868-839-11
12	22	32.8	13	4	US-09-811-672-19
13	21	31.3	13	1	US-08-068-947-19
14	20.5	30.6	13	3	US-09-040-216-8
15	20	29.9	13	1	US-08-068-947-21
16	20	29.9	13	3	US-08-630-916A-6
17	20	29.9	13	3	US-08-630-916A-6
18	20	29.9	13	4	US-08-630-916A-6
19	20	29.9	13	4	US-09-297-369A-7
20	20	29.9	13	4	US-09-635-872A-52
21	20	29.9	13	4	US-09-636-077A-52
22	20	29.9	13	4	US-09-543-608A-29
23	20	29.9	13	4	US-09-543-608A-30
24	20	29.9	13	4	US-09-543-608A-31
25	20	29.9	13	4	US-09-636-060C-52
26	20	29.9	13	4	US-09-986-552-52
27	19	28.4	13	1	US-08-068-947-8

28	19	28.4	13	1	US-08-068-947-24	Sequence 24, Appl
29	19	28.4	13	1	US-08-469-005A-13	Sequence 13, Appl
30	19	28.4	13	1	US-08-188-426-7	Sequence 7, Appl
31	19	28.4	13	1	US-08-471-033-11	Sequence 11, Appl
32	19	28.4	13	2	US-08-471-044-11	Sequence 11, Appl
33	19	28.4	13	2	US-08-463-483A-11	Sequence 11, Appl
34	19	28.4	13	2	US-08-471-046A-11	Sequence 11, Appl
35	19	28.4	13	2	US-08-470-566B-11	Sequence 11, Appl
36	19	28.4	13	2	US-08-469-007-7	Sequence 7, Appl
37	19	28.4	13	2	US-08-838-219B-11	Sequence 11, Appl
38	19	28.4	13	2	US-08-469-009-7	Sequence 11, Appl
39	19	28.4	13	2	US-08-630-916A-58	Sequence 58, Appl
40	19	28.4	13	3	US-08-881-094-42	Sequence 42, Appl
41	19	28.4	13	3	US-09-300-529-11	Sequence 11, Appl
42	19	28.4	13	3	US-09-233-336A-11	Sequence 11, Appl
43	19	28.4	13	3	US-09-233-752A-11	Sequence 11, Appl
44	19	28.4	13	3	US-09-094-557-5	Sequence 5, Appl
45	19	28.4	13	3		

## ALIGNMENTS

```

RESULT 1
US-09-675-922-1
; Sequence 1, Application US/09675922
; Patent No. 6468731
; GENERAL INFORMATION:
; APPLICANT: Hubbell A., Jeffrey
; APPLICANT: Schense C., Jason
; APPLICANT: Sakiyama E., Shelley
; TITLE OF INVENTION: Enzyme-Mediated Modification of Fibrin for Tissue
; FILE REFERENCE: ETH 107 DIV
; CURRENT APPLICATION NUMBER: US/09/675,922
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Factor XIIIa at
; OTHER INFORMATION: N-terminus and modified peptide from the
; NAME/KEY: MOD RES
; LOCATION: (1)
; OTHER INFORMATION: dansyl leucine
US-09-675-922-1

Query Match          53.7% Score 36; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NOEQVSP 7
      |||||
Db      2 NOEQVSP 8

RESULT 2
US-08-174-467-6
; Sequence 6, Application US/08174467
; Patent No. 5451514
; GENERAL INFORMATION:
; APPLICANT: BOUDET, ALAIN M.
; APPLICANT: INZE, DIRK G.
; APPLICANT: SCHUCH, WOLFGANG W.
; TITLE OF INVENTION: MODIFICATION OF LIGNIN SYNTHESIS IN
; TITLE OF INVENTION: PLANTS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESS: CUSHMAN, DARBY & CUSHMAN

```

STREET: 1100 NEW YORK AVENUE, N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/174,467  
FILING DATE: 28-DEC-1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/874,166  
FILING DATE: 27-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, PAUL N.  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: 95563/PS36321/US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-174-467-6

Query Match 35.8%; Score 24; DB 1; Length 13;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 VSPYT 9  
Db 5 LSPYT 9

RESULT 3  
US-08-452-071-6  
Sequence 6, Application US/08452071  
Patent No. 6066780  
GENERAL INFORMATION:  
APPLICANT: BOUBET, ALAIN M.  
APPLICANT: INZE, DIRK G.  
APPLICANT: SCHUCH, WOLFGANG W.  
TITLE OF INVENTION: MODIFICATION OF LIGNIN SYNTHESIS IN  
TITLE OF INVENTION: PLANTS  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
STREET: 1100 NEW YORK AVENUE, N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/452,071  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/874,166  
FILING DATE: 27-APR-1992  
ATTORNEY/AGENT INFORMATION:

NAME: KOKULIS, PAUL N.  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: 95563/PS36321/US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-452-071-6

Query Match 35.8%; Score 24; DB 3; Length 13;  
Best Local Similarity 80.0%; Pred. No. 2.8e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 VSPYT 9  
Db 5 LSPYT 9

RESULT 4  
US-08-620-213-3  
Sequence 3, Application US/08620213  
Patent No. 5677297  
GENERAL INFORMATION:  
APPLICANT: WALDECK, Harald  
APPLICANT: HOELTJE, Dagmar  
APPLICANT: MESSINGER, Josef  
APPLICANT: ANTEL, Jochem  
APPLICANT: WURL, Michael  
APPLICANT: THORMAEHLER, Dirk  
TITLE OF INVENTION: BENZAZEPINE-, BENZOAZEPINE- AND  
TITLE OF INVENTION: BENZOTHAZEPINE-N-ACETIC ACID DERIVATIVES, PROCESS FOR THEIR  
TITLE OF INVENTION: PREPARATION AND PHARMACEUTICAL COMPOSITIONS CONTAINING THEM  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
STREET: 1200 G Street, N.W., Suite 700  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/620,213  
FILING DATE:  
CLASSIFICATION: 544  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 195 10 566.4  
FILING DATE: 23-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 181/42626  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-8800  
TELEFAX: (202) 628-8844  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-620-213-3

Query Match 34.3%; Score 23; DB 1; Length 13;  
Best Local Similarity 62.5%; Pred. No. 4.3e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 EQVSPYTLK 10  
Db 5 EHVPYGL 12

RESULT 5

US-08-542-686-1  
; Sequence 1, Application US/08542686  
; Patent No. 5824487  
; GENERAL INFORMATION:  
; APPLICANT: Ravetch, Jeffrey V.  
; APPLICANT: Kurosaki, Tomohiro  
; TITLE OF INVENTION: METHOD FOR SCREENING FOR TARGETS FOR  
; TITLE OF INVENTION: ANTI-INFLAMMATORY OR ANTI-ALLERGIC AGENTS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/542,686  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/052,269  
; FILING DATE: 23-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John J.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 42113/JPW/AKC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 977-9550  
; TELEFAX: (212) 664-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
US-08-542-686-1

Query Match 34.3%; Score 23; DB 2; Length 13;  
Best Local Similarity 50.0%; Pred. No. 4.3e+02;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 EQVSPYTLK 12  
Db 2 ENTITSLK 11

RESULT 6  
US-08-332-562A-69  
; Sequence 69, Application US/08332562A  
; Patent No. 5985599  
; GENERAL INFORMATION:  
; APPLICANT: MCKENZIE, Ian F.C.

APPLICANT: HOGARTH, Mark P.  
; APPLICANT: HIBBS, Margaret L.  
; APPLICANT: SCOTT, Bernadette M.  
; APPLICANT: BONADONNA, Lisa  
; APPLICANT: HULETT, Mark D.  
; TITLE OF INVENTION: FC RECEPTOR FOR IMMUNOGLOBULIN  
; NUMBER OF SEQUENCES: 136  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/332,562A  
; FILING DATE: 31-OCT-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/896,457  
; FILING DATE: 27-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 54270/119/GRHA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 69:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-332-562A-69

Query Match 34.3%; Score 23; DB 2; Length 13;  
Best Local Similarity 50.0%; Pred. No. 4.3e+02;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 EQVSPYTLK 12  
Db 1 ENTITSLK 10

RESULT 7  
US-08-068-947-14  
; Sequence 14, Application US/08068947  
; Patent No. 5470753  
; GENERAL INFORMATION:  
; APPLICANT: Sepelev, Nikolai  
; APPLICANT: Issekova, Olga  
; APPLICANT: Krachak, Viktor  
; APPLICANT: Lebl, Michel  
; TITLE OF INVENTION: PEPTIDE SEQUENCING USING MASS  
; TITLE OF INVENTION: SPECTROMETRY  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; SOFTWARE: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/068,947  
FILING DATE: 19930528  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock Esq., S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 7156-040-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-068-947-14

Query Match 32.8%; Score 22; DB 1; Length 13;  
Best Local Similarity 42.9%; Pred. No. 6.7e+02;  
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NQGVSP 7  
|:|:|  
Db 1 NADQIQ 7

RESULT 8  
US-08-068-947-15  
Sequence 15, Application US/08068947  
Patent No. 5470753  
GENERAL INFORMATION:  
APPLICANT: Sepetov, Nikolai  
APPLICANT: Lesakova, Olga  
APPLICANT: Krchnak, Viktor  
APPLICANT: Iedl, Michal  
TITLE OF INVENTION: PEPTIDE SEQUENCING USING MASS  
NUMBER OF SEQUENCES: 47  
CURRENT APPLICATION DATA:  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/068,947  
FILING DATE: 19930528  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock Esq., S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 7156-040-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: unknown

TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-068-947-15

Query Match 32.8%; Score 22; DB 1; Length 13;  
Best Local Similarity 50.0%; Pred. No. 6.7e+02;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVSP 7  
|:|:|  
Db 2 QDQIQ 7

RESULT 9  
US-08-750-419A-22  
Sequence 22, Application US/08750419A  
Patent No. 6008340  
GENERAL INFORMATION:  
APPLICANT: BALL, TANJA  
APPLICANT: VITALA, SUSANNE  
APPLICANT: SPERR, WOLFGANG  
APPLICANT: VALENT, PETER  
APPLICANT: SUSANI, MARKUS  
APPLICANT: KRAFT, DIETRICH  
APPLICANT: LAFFER, SYLVIA  
TITLE OF INVENTION: RECOMBINANT ALLERGEN, FRAGMENTS THEREOF, VECTORS AND HOSTS  
TITLE OF INVENTION: CORRESPONDING RECOMBINANT DNA MOLECULES, DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: CONTAINING THE DNA MOLECULES, DIAGNOSTIC AND THERAPEUTIC  
TITLE OF INVENTION: USES OF SAID ALLERGENS AND FRAGMENTS  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH  
STREET: PO BOX 747  
CITY: FALLS CHURCH  
STATE: VA  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/750,419A  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: MURPHY JR, GERALD M  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1614-175  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8050  
TELEFAX: (703) 205-8050  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-750-419A-22

Query Match 32.8%; Score 22; DB 3; Length 13;  
Best Local Similarity 50.0%; Pred. No. 6.7e+02;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QGVSPY 8  
|:|:|  
Db 1 EPIAPY 6

RESULT 10  
US-09-468-578-11

```
; Sequence 11, Application US/09468578
; Patent No. 639329
; GENERAL INFORMATION:
; APPLICANT: Wang, Huaming
; APPLICANT: Bodie, Elizabeth A.
; TITLE OF INVENTION: Phenol Oxidizing Enzymes
; FILE REFERENCE: GC561-3
; CURRENT APPLICATION NUMBER: US/09/468,578
; CURRENT FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: US 09/220,871
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 09/338,723
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Stachybotrys chartarum
US-09-468-578-11
```

```
Query Match          32.8%; Score 22; DB 4; Length 13;
Best Local Similarity 80.0%; Pred. No. 6.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      4 QVSPY 8
      |||||
Db      3 QVMPY 7
```

```
RESULT 11
US-09-868-839-11
; Sequence 11, Application US/09868839
; Patent No. 6509307
; GENERAL INFORMATION:
; APPLICANT: UNILEVER N.V. et al.
; TITLE OF INVENTION: DETERGENT COMPOSITIONS COMPRISING PHENOL OXIDIZING
; FILE REFERENCE: C7516(V)
; CURRENT APPLICATION NUMBER: US/09/868,839
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: US 09/220,871
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 09/338,723
; PRIOR FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Stachybotrys chartarum
US-09-868-839-11
```

```
Query Match          32.8%; Score 22; DB 4; Length 13;
Best Local Similarity 80.0%; Pred. No. 6.7e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      4 QVSPY 8
      |||||
Db      3 QVMPY 7
```

```
RESULT 12
US-09-811-672-19
; Sequence 19, Application US/09811672
; Patent No. 6559120
; GENERAL INFORMATION:
; APPLICANT: BALL, Tanja
; APPLICANT: VRTALA, Susanne
; APPLICANT: SPERR, Wolfgang
; APPLICANT: VALENT, Peter
; APPLICANT: SUSANI, Markus
; APPLICANT: KRAFT, Dietrich
```

```
; APPLICANT: VALENTA, Rudolf
; APPLICANT: LAFFER, Sylvia
; TITLE OF INVENTION: RECOMBINANT ALLERGEN, FRAGMENTS THEREOF, CORRESPONDING RECOMBINANT
; TITLE OF INVENTION: MOLECULES, VECTORS AND HOSTS CONTAINING THE DNA MOLECULES, DIAGN
; FILE REFERENCE: 1614-0247P
; CURRENT APPLICATION NUMBER: US/09/811,672
; CURRENT FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Timothy Grass
US-09-811-672-19
```

```
Query Match          32.8%; Score 22; DB 4; Length 13;
Best Local Similarity 50.0%; Pred. No. 6.7e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      3 EQVSPY 8
      |||||
Db      1 EPIAPY 6
```

```
RESULT 13
US-08-068-947-19
; Sequence 19, Application US/08068947
; Patent No. 5470753
; GENERAL INFORMATION:
; APPLICANT: Sepelov, Nikolai
; APPLICANT: Issakova, Olga
; APPLICANT: Kirchak, Viktor
; APPLICANT: Leb1, Michael
; TITLE OF INVENTION: PEPTIDE SEQUENCING USING MASS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Penile & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,947
; FILING DATE: 19930528
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock Esq., S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7156-040-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-068-947-19
```

```
Query Match          31.3%; Score 21; DB 1; Length 13;
Best Local Similarity 42.9%; Pred. No. 1.1e+03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

QY 1 NOEOWSP 7  
| : | : |  
Db 1 NEGQIQP 7

## RESULT 14

US-09-040-216-8  
; Sequence 8, Application US/09040216  
; Patent No. 6030942  
; GENERAL INFORMATION:  
; APPLICANT: COOPERMAN, ET AL., BARRY  
; TITLE OF INVENTION: PEPTIDES, PEPTIDE ANALOGS, PEPTIDOMIMETICS, AND OTHER  
; TITLE OF INVENTION: SMALL MOLECULES USEFUL FOR INHIBITING THE ACTIVITY OF  
; FILE REFERENCE: 9596-63U1  
; CURRENT APPLICATION NUMBER: US/09/040,216  
; CURRENT FILING DATE: 1998-03-17  
; EARLIER APPLICATION NUMBER: 08/919,748  
; EARLIER FILING DATE: 1997-08-28  
; EARLIER APPLICATION NUMBER: 60/025,146  
; EARLIER FILING DATE: 1996-08-30  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
; FEATURE:  
; OTHER INFORMATION: alpha 13 helix region of E. coli R1  
US-09-040-216-8

Query Match 30.6%; Score 20.5; DB 3; Length 13;  
Best Local Similarity 85.7%; Pred. No. 1.3e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 8 YT-LKKG 13  
| : | : | : | : |  
Db 5 YTRLKKG 11

RESULT 15  
US-08-068-947-21  
; Sequence 21, Application US/08068947  
; Patent No. 5470753  
; GENERAL INFORMATION:  
; APPLICANT: Seetov, Nikolai  
; APPLICANT: Issakova, Olga  
; APPLICANT: Kirchnak, Viktor  
; APPLICANT: Lebl, Michael  
; TITLE OF INVENTION: PEPTIDE SEQUENCING USING MASS  
; TITLE OF INVENTION: SPECTROMETRY  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/068,947  
; FILING DATE: 19930528  
; CLASSIFICATION: 436  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock Esq., S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 7156-040-999

## TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 13 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-068-947-21

Query Match 29.9%; Score 20; DB 1; Length 13;  
Best Local Similarity 60.0%; Pred. No. 1.6e+03;  
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 QVSPY 8  
| : | : | : |  
Db 9 QVTPF 13

Search completed: October 1, 2004, 13:50:30  
Job time : 18 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: October 1, 2004, 13:46:59 ; Search time 45 Seconds  
(without alignments)  
92.964 Million cell updates/sec

Title: US-09-674-616A-2  
Perfect score: 67  
Sequence: 1 NOEQVSPYTLKG 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 11839

Minimum DB seq length: 13  
Maximum DB seq length: 13

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubppaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubppaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	36	53.7	13	14	US-10-323-046-22
2	24	35.8	13	10	US-09-932-613-156
3	24	35.8	13	10	US-09-932-322-156
4	24	35.8	13	15	US-10-361-208-64
5	24	35.8	13	15	US-10-361-208-197
6	23	34.3	13	14	US-10-226-007-782
7	23	34.3	13	14	US-10-226-007-794
8	23	34.3	13	14	US-10-226-007-805
9	23	34.3	13	14	US-10-226-007-815
10	23	34.3	13	14	US-10-226-007-824
11	23	34.3	13	14	US-10-226-007-832
12	22	32.8	13	15	US-10-412-897-20
13	22	32.8	13	9	US-09-338-723A-7
14	22	32.8	13	9	US-09-811-672-19
15	22	32.8	13	9	US-09-946-175-8

16	22	32.8	13	9	US-09-791-378-594	Sequence 594, App
17	22	32.8	13	10	US-09-791-393-156	Sequence 156, App
18	22	32.8	13	10	US-09-791-389-156	Sequence 156, App
19	22	32.8	13	12	US-09-791-377-594	Sequence 594, App
20	22	32.8	13	13	US-10-080-210-11	Sequence 11, App
21	22	32.8	13	14	US-10-153-244-78	Sequence 78, App
22	22	32.8	13	14	US-10-153-244-141	Sequence 141, App
23	22	32.8	13	14	US-10-153-244-191	Sequence 191, App
24	22	32.8	13	14	US-10-153-244-241	Sequence 241, App
25	22	32.8	13	14	US-10-161-660-22	Sequence 22, App
26	22	32.8	13	15	US-10-285-394-162	Sequence 162, App
27	22	32.8	13	16	US-10-311-508-19	Sequence 19, App
28	21	31.3	13	10	US-09-993-180-13	Sequence 13, App
29	21	31.3	13	14	US-10-104-943-12	Sequence 12, App
30	21	31.3	13	14	US-10-104-943-57	Sequence 57, App
31	21	31.3	13	14	US-10-226-007-1016	Sequence 1016, App
32	21	31.3	13	14	US-10-226-007-1029	Sequence 1029, App
33	21	31.3	13	14	US-10-226-007-1042	Sequence 1042, App
34	21	31.3	13	14	US-10-226-007-1055	Sequence 1055, App
35	21	31.3	13	14	US-10-226-007-1068	Sequence 1068, App
36	21	31.3	13	14	US-10-224-999A-1786	Sequence 23, App
37	21	31.3	13	14	US-10-224-999A-1787	Sequence 1786, App
38	21	31.3	13	14	US-10-224-999A-1787	Sequence 1787, App
39	21	31.3	13	14	US-10-224-999A-1788	Sequence 1788, App
40	21	31.3	13	14	US-10-224-999A-1789	Sequence 1789, App
41	21	31.3	13	14	US-10-224-999A-1790	Sequence 1790, App
42	21	31.3	13	15	US-10-298-461-14	Sequence 14, App
43	21	31.3	13	15	US-10-285-394-129	Sequence 129, App
44	21	31.3	13	16	US-10-652-407-8	Sequence 8, App
45	20	29.9	13	9	US-09-895-072-52	Sequence 52, App

## ALIGNMENTS

RESULT 1  
US-10-323-046-22  
; Sequence 22, Application US/10323046  
; Publication No. US20030187232A1  
; GENERAL INFORMATION:  
; APPLICANT: Hubbell, Jeffrey A  
; APPLICANT: Schemse, Jason C  
; APPLICANT: Sakiyama-Elbert, Shelly E  
; TITLE OF INVENTION: Growth Factor Modified Protein Matrices for Tissue  
; FILE REFERENCE: ETH 107 CIP (2)  
; CURRENT APPLICATION NUMBER: US/10/323,046  
; CURRENT FILING DATE: 2002-12-17  
; PRIOR APPLICATION NUMBER: 09/141,153  
; PRIOR FILING DATE: 1998-08-27  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 22  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Peptide  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)  
; OTHER INFORMATION: dansyl Leucine  
US-10-323-046-22

Query Match 53.7%: Score 36; DB 14; length 13;  
Best Local Similarity 100.0%; Pred.No. 7.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOEQVSP 7  
|||||  
Db 2 NOEQVSP 8

```
RESULT 2
US-09-932-613-156
; Sequence 156, Application US/09932613
; Publication No. US20030091565A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; APPLICANT: Belter, James P.
; APPLICANT: Poter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
; FILE REFERENCE: DYX-025.1 PCT; DYX-025.1 US
; CURRENT APPLICATION NUMBER: US/09/932,613
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 156
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Blys binding polypeptide
US-09-932-613-156
Query Match      35.8%; Score 24; DB 10; Length 13;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      4 QVSPYTLK 12
      |||||
      2 QVDEPTGLK 10

RESULT 3
US-09-932-322-156
; Sequence 156, Application US/09932322
; Publication No. US20030194743A1
; GENERAL INFORMATION:
; APPLICANT: Dyax Corp.
; APPLICANT: Belter, James P.
; APPLICANT: Poter, M. Daniel
; APPLICANT: Fleming, Tony J.
; APPLICANT: Ladner, Robert Charles
; TITLE OF INVENTION: BINDING POLYPEPTIDES FOR B LYMPHOCYTE STIMULATOR PROTEIN (Blys)
; FILE REFERENCE: DYX-018.1 PCT; DYX-018.1 US
; CURRENT APPLICATION NUMBER: US/09/932,322
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 458
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 156
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Blys binding polypeptide
US-09-932-322-156
Query Match      35.8%; Score 24; DB 10; Length 13;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      4 QVSPYTLK 12
      |||||
      2 QVDEPTGLK 10

RESULT 4
US-10-361-208-64
; Sequence 64, Application US/10361208
; Publication No. US20040009167A1
; GENERAL INFORMATION:
; APPLICANT: Rider, Todd H.
; TITLE OF INVENTION: ANTI-PATHOGEN TREATMENTS
```

```
; FILE REFERENCE: 0050.2041-003
; CURRENT APPLICATION NUMBER: US/10/361,208
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,359
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/355,022
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/432,386
; PRIOR FILING DATE: 2002-12-10
; NUMBER OF SEQ ID NOS: 473
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 13
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: partial Hdj-1 sequence
; NAME/KEY: NON_CONS
; LOCATION: 7-8
US-10-361-208-64
Query Match      35.8%; Score 24; DB 15; Length 13;
Best Local Similarity 83.3%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 QEOVSP 7
      |||||
      7 QEOVLP 12

RESULT 5
US-10-361-208-197
; Sequence 197, Application US/10361208
; Publication No. US20040009167A1
; GENERAL INFORMATION:
; APPLICANT: Rider, Todd H.
; TITLE OF INVENTION: ANTI-PATHOGEN TREATMENTS
; FILE REFERENCE: 0050.2041-003
; CURRENT APPLICATION NUMBER: US/10/361,208
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 60/355,359
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/355,022
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/432,386
; PRIOR FILING DATE: 2002-12-10
; NUMBER OF SEQ ID NOS: 473
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 197
; LENGTH: 13
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; OTHER INFORMATION: partial Hdj-1 sequence
; NAME/KEY: NON_CONS
; LOCATION: 7-8
US-10-361-208-197
Query Match      35.8%; Score 24; DB 15; Length 13;
Best Local Similarity 83.3%; Pred. No. 1.3e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 QEOVSP 7
      |||||
      7 QEOVLP 12

RESULT 6
US-10-226-007-782
; Sequence 782, Application US/10226007
; Publication No. US20030105277A1
```



```
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 782
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Human herpesvirus 4
US-10-226-007-782
```

```
Query Match          34.3%; Score 23; DB 14; Length 13;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 NCEQVSPY 8
        |:.|:|
Db       1 NEEPPPPY 8
```

```
RESULT 7
US-10-226-007-794
; Sequence 794, Application US/10226007
; Publication No. US20030105277A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 794
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Human herpesvirus 4
US-10-226-007-794
```

```
Query Match          34.3%; Score 23; DB 14; Length 13;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 NCEQVSPY 8
        |:.|:|
Db       2 NEEPPPPY 9
```

```
RESULT 8
US-10-226-007-805
; Sequence 805, Application US/10226007
; Publication No. US20030105277A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
```

```
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 805
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Human herpesvirus 4
US-10-226-007-805
```

```
Query Match          34.3%; Score 23; DB 14; Length 13;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 NCEQVSPY 8
        |:.|:|
Db       3 NEEPPPPY 10
```

```
RESULT 9
US-10-226-007-815
; Sequence 815, Application US/10226007
; Publication No. US20030105277A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 815
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Human herpesvirus 4
US-10-226-007-815
```

```
Query Match          34.3%; Score 23; DB 14; Length 13;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      1 NCEQVSPY 8
        |:.|:|
Db       4 NEEPPPPY 11
```

```
RESULT 10
US-10-226-007-824
; Sequence 824, Application US/10226007
; Publication No. US20030105277A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 824
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Human herpesvirus 4
US-10-226-007-824
```

Query Match 34.3%; Score 23; DB 14; Length 13;  
Best Local Similarity 50.0%; Pred. No. 2e+03;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NOEQVSPY 8  
|:|:|  
Db 5 NEEPPPPY 12

## RESULT 11

US-10-226-007-832  
; Sequence 832, Application US/10226007  
; Publication No. US20030105277A1  
; GENERAL INFORMATION:  
; APPLICANT: Myriad Genetics, Inc.  
; APPLICANT: Morham, Scott  
; APPLICANT: Zavit, Kenton  
; APPLICANT: Hobden, Adrian  
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection  
; FILE REFERENCE: 5005.01  
; CURRENT FILING DATE: 2002-11-15  
; PRIOR APPLICATION NUMBER: US 60/313,883  
; PRIOR FILING DATE: 2001-08-21  
; NUMBER OF SEQ ID NOS: 1673  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 832  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Human herpesvirus 4  
US-10-226-007-832

Query Match 34.3%; Score 23; DB 14; Length 13;  
Best Local Similarity 50.0%; Pred. No. 2e+03;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 NOEQVSPY 8  
|:|:|  
Db 6 NEEPPPPY 13

## RESULT 12

US-10-412-897-20  
; Sequence 20, Application US/10412897  
; Publication No. US20030220224A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES ENCODING THE HUMAN CITRON KINASE  
; FILE REFERENCE: D0193 NP  
; CURRENT APPLICATION NUMBER: US/10/412,897  
; CURRENT FILING DATE: 2003-04-11  
; PRIOR APPLICATION NUMBER: U.S. 60/372,745  
; PRIOR FILING DATE: 2002-04-12  
; NUMBER OF SEQ ID NOS: 77  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 20  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-412-897-20

Query Match 34.3%; Score 23; DB 15; Length 13;  
Best Local Similarity 100.0%; Pred. No. 2e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OEQVS 6  
|:|:|  
Db 2 OEQVS 6

## RESULT 13

## US-09-338-723A-7

; Sequence 7, Application US/09338723A  
; Patent No. US20020019038A1  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Wang  
; TITLE OF INVENTION: Phenol Oxidizing Enzymes  
; FILE REFERENCE: GC561-2  
; CURRENT APPLICATION NUMBER: US/09/338,723A  
; CURRENT FILING DATE: 1999-06-23  
; PRIOR APPLICATION NUMBER: 09/220,871  
; PRIOR FILING DATE: 1998-12-23  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Stachybotrys chartarum  
US-09-338-723A-7

Query Match 32.8%; Score 22; DB 9; Length 13;  
Best Local Similarity 80.0%; Pred. No. 3.2e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 QVSPY 8  
|:|:|  
Db 3 QVMPY 7

## RESULT 14

US-09-811-672-19  
; Sequence 19, Application US/09811672  
; Patent No. US20020052490A1  
; GENERAL INFORMATION:  
; APPLICANT: BALL, Tania  
; APPLICANT: VITALA, Susanne  
; APPLICANT: SPEER, Wolfgang  
; APPLICANT: VALENT, Peter  
; APPLICANT: SUSANI, Markus  
; APPLICANT: KRAFT, Dietrich  
; APPLICANT: VALENTA, Rudolf  
; APPLICANT: LAFFER, Sylvia  
; TITLE OF INVENTION: RECOMBINANT ALLERGEN, FRAGMENTS THEREOF, CORRESPONDING RECOMBINANT  
; TITLE OF INVENTION: MOLECULES, VECTORS AND HOSTS CONTAINING THE DNA MOLECULES, DIAGN  
; FILE REFERENCE: 1614-0247P  
; CURRENT APPLICATION NUMBER: US/09/811,672  
; CURRENT FILING DATE: 2001-03-20  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Timothy Grass  
US-09-811-672-19

Query Match 32.8%; Score 22; DB 9; Length 13;  
Best Local Similarity 50.0%; Pred. No. 3.2e+03;  
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 EQVSPY 8  
|:|:|  
Db 1 EPIARY 6

## RESULT 15

US-09-946-175-8  
; Sequence 8, Application US/09946175  
; Patent No. US20020106671A1  
; GENERAL INFORMATION:  
; APPLICANT: Warner-Lambert  
; TITLE OF INVENTION: IDENTIFICATION OF A CAPACITATIVE CALCIUM  
; FILE REFERENCE: CHANNEL IN ANTIGEN PRESENTING CELLS AND USES THEREOF  
US-09-946-175-8

```

; CURRENT APPLICATION NUMBER: US/09/946,175
; CURRENT FILING DATE: 2001-09-05
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentm Ver. 2.1
; SEQ ID NO 8
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-946-175-8

```

```

Query Match          32.8%; Score 22; DB 9; Length 13;
Best Local Similarity 33.3%; Pred. No. 3.2e+03;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 5 VSPYTLKG 13
   :||::|
Db 4 LSPYVMMIG 12

```

```

Search completed: October 1, 2004, 13:49:24
Job time : 46 secs

```

Blank Sheet

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: October 1, 2004, 13:46:58 ; Search time 51 Seconds  
(without alignments)  
72.022 Million cell updates/sec

Title: US-09-674-616A-2  
Perfect score: 67  
Sequence: 1 NQEOVSPYTLKG 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 15593

Minimum DB seq length: 13  
Maximum DB seq length: 13

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	100.0	13	3	AAV67835
2	64	95.5	13	3	AAV67837
3	61	91.0	13	3	AAV67836
4	61	91.0	13	3	AAV67842
5	59	88.1	13	3	AAV67841
6	59	88.1	13	3	AAV67840
7	46	68.7	13	3	AAV67844
8	36	53.7	13	3	AAV67844
9	36	53.7	13	6	ABG73917
10	27	40.3	13	2	AAV20946
11	26	38.8	13	2	AAV20420
12	25	37.3	13	2	AAV20420
13	25	37.3	13	2	AAV20420
14	25	37.3	13	2	AAV20420
15	24	35.8	13	2	AAV20420
16	24	35.8	13	2	AAV20420
17	24	35.8	13	2	AAV20420
18	24	35.8	13	2	AAV20420
19	24	35.8	13	2	AAV20420
20	23	34.3	13	2	AAV20420
21	23	34.3	13	2	AAV20420
22	23	34.3	13	2	AAV20420
23	23	34.3	13	2	AAV20420
24	23	34.3	13	2	AAV20420
25	23	34.3	13	2	AAV20420

26	23	34.3	13	4	AAV67835	AAV67835
27	23	34.3	13	5	AAV67835	AAV67835
28	23	34.3	13	6	AAV67835	AAV67835
29	23	34.3	13	6	AAV67835	AAV67835
30	23	34.3	13	6	AAV67835	AAV67835
31	23	34.3	13	6	AAV67835	AAV67835
32	23	34.3	13	6	AAV67835	AAV67835
33	23	34.3	13	6	AAV67835	AAV67835
34	23	34.3	13	6	AAV67835	AAV67835
35	23	34.3	13	6	AAV67835	AAV67835
36	23	34.3	13	6	AAV67835	AAV67835
37	23	34.3	13	6	AAV67835	AAV67835
38	23	34.3	13	6	AAV67835	AAV67835
39	23	34.3	13	6	AAV67835	AAV67835
40	23	34.3	13	6	AAV67835	AAV67835
41	23	34.3	13	6	AAV67835	AAV67835
42	23	34.3	13	6	AAV67835	AAV67835
43	23	34.3	13	6	AAV67835	AAV67835
44	23	34.3	13	6	AAV67835	AAV67835
45	23	34.3	13	6	AAV67835	AAV67835

## ALIGNMENTS

RESULT 1	AAV67835	standard; peptide; 13 AA.
ID	AAV67835	
AC	AAV67835	
DT	25-APR-2000	(first entry)
XX		
DE	Alpha-2 antiplasmin enzyme derived peptide #2.	
XX		
KW	Alpha-2 antiplasmin; fibrin stabilising enzyme; Factor XIIIa; cancer; embolism; thrombosis; inflammation; atherosclerosis; diagnosis.	
OS	Synthetic.	
XX		
FT	Key	Location/Qualifiers
FT	Modified-site	1 /note="Asn optionally forms a peptide bond with C-terminal Gly to form a cyclic peptide"
FT	Misc-difference	5..13 /note="Optionally D form residues"
FT	Modified-site	13 /note="Gly optionally forms a peptide bond with N-terminal Asn to form a cyclic peptide"
FT		
XX		
PN	WO9960018-A1.	
XX		
PD	25-NOV-1999.	
XX		
PF	14-MAY-1999; 99WO-GB001550.	
XX		
PR	15-MAY-1998; 98EP-00303872.	
XX		
PA	(AMSH) NYCOMED AMERSHAM PLC.	
XX		
PI	Storey AE, Mendisabal M, Champion S, Gibson A, Guilbert B; Wilson IA, Knox P;	
XX		
DR	WPI; 2000-126380/11.	
XX		
PT	Novel synthetic analogues for diagnosis of thrombosis, embolism, atherosclerosis, inflammation or cancer.	
XX		
PS	Example 1; Page 16; 46pp; English.	
XX		
CC	This sequence represents an alpha-2 antiplasmin derived peptide. Alpha-2 antiplasmin is a subunit for the fibrin stabilising enzyme Factor XIIIa. Synthetic analogues of lysine and glutamine labelled with a	

CC detectable moiety can also act as substrates for Factor XIIIa. The  
 CC invention relates to compounds which can be used in the diagnosis of  
 CC thrombosis or embolism and also for diagnosing atherosclerosis,  
 CC inflammation or cancer. The peptide forms part of the compound

XX Sequence 13 AA;

Query Match 100.0%; Score 67; DB 3; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOEQVSPYTLKKG 13  
 |||:|||||  
 Db 1 NOEQVSPYTLKKG 13

RESULT 2

AAV67837  
 ID AAV67837 standard; peptide; 13 AA.

XX AAV67837;

XX 25-APR-2000 (first entry)

DE Alpha-2 antipiasmin enzyme derived peptide #4.

KM Alpha-2 antipiasmin; fibrin stabilising enzyme; Factor XIIIa; cancer;  
 KM embolism; thrombosis; inflammation; atherosclerosis; diagnosis.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

XX W09960018-A1.

XX 25-NOV-1999.

XX 14-MAY-1999; 99WO-GB001550.

XX 15-MAY-1998; 98EP-00303872.

XX (AMSH ) NYCOMED AMERSHAM PLC.

XX Storey AE, Mendizabal M, Champion S, Gibson A, Gullbert B;  
 PI Wilson IA, Knox P;

XX WPI: 2000-126380/11.

XX Novel synthetic analogues for diagnosis of thrombosis, embolism,  
 PT atherosclerosis, inflammation or cancer.

XX Example 2; Page 16; 46pp; English.

XX This sequence represents an alpha-2 antipiasmin derived peptide. Alpha-2  
 CC antipiasmin is a substrate for the fibrin stabilising enzyme Factor  
 CC XIIIa. Synthetic analogues of lysine and glutamine labelled with a  
 CC detectable moiety can also act as substrates for Factor XIIIa. The  
 CC invention relates to compounds which can be used in the diagnosis of  
 CC thrombosis or embolism and also for diagnosing atherosclerosis,  
 CC inflammation or cancer. The peptide forms part of the compound

XX Sequence 13 AA;

Query Match 95.5%; Score 64; DB 3; Length 13;  
 Best Local Similarity 92.3%; Pred. No. 4.3e-05;  
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOEQVSPYTLKKG 13  
 |||:|||||  
 Db 1 NOEQVSPYTLKKG 13

RESULT 3

AAV67836  
 ID AAV67836 standard; peptide; 13 AA.

XX AAV67836;

XX 25-APR-2000 (first entry)

DE Alpha-2 antipiasmin enzyme derived peptide #3.

KM Alpha-2 antipiasmin; fibrin stabilising enzyme; Factor XIIIa; cancer;  
 KM embolism; thrombosis; inflammation; atherosclerosis; diagnosis.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

XX W09960018-A1.

XX 25-NOV-1999.

XX 14-MAY-1999; 99WO-GB001550.

XX 15-MAY-1998; 98EP-00303872.

XX (AMSH ) NYCOMED AMERSHAM PLC.

XX Storey AE, Mendizabal M, Champion S, Gibson A, Gullbert B;  
 PI Wilson IA, Knox P;

XX WPI: 2000-126380/11.

XX Novel synthetic analogues for diagnosis of thrombosis, embolism,  
 PT atherosclerosis, inflammation or cancer.

XX Example 2; Page 16; 46pp; English.

XX This sequence represents an alpha-2 antipiasmin derived peptide. Alpha-2  
 CC antipiasmin is a substrate for the fibrin stabilising enzyme Factor  
 CC XIIIa. Synthetic analogues of lysine and glutamine labelled with a  
 CC detectable moiety can also act as substrates for Factor XIIIa. The  
 CC invention relates to compounds which can be used in the diagnosis of  
 CC thrombosis or embolism and also for diagnosing atherosclerosis,  
 CC inflammation or cancer. The peptide forms part of the compound

XX Sequence 13 AA;

Query Match 91.0%; Score 61; DB 3; Length 13;  
 Best Local Similarity 92.3%; Pred. No. 0.00016;  
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NOEQVSPYTLKKG 13  
 |||:|||||  
 Db 1 NOEQVSPYTLKKG 13

RESULT 4

AAV67842  
 ID AAV67842 standard; peptide; 13 AA.

XX AAV67842;

XX 25-APR-2000 (first entry)

DE Alpha-2 antipiasmin enzyme derived peptide #9.

KM Alpha-2 antipiasmin; fibrin stabilising enzyme; Factor XIIIa; cancer;  
 KM embolism; thrombosis; inflammation; atherosclerosis; diagnosis.

XX Synthetic.

```

XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "N-terminal acetyl"
FT Misc-difference 8
FT /note= "D-form residue"
FT Misc-difference 12
FT /note= "D-form residue"
FT Modified-site 13
FT /label= bAla
XX WO960018-A1.
XX PD
XX 25-NOV-1999.
XX PF 14-MAY-1999; 99WO-GB001550.
XX PR 15-MAY-1998; 98EP-00303872.
XX (AMSH ) NYCOMED AMERSHAM PLC.
XX PA
XX Storey AE, Mendizabal M, Champion S, Gibson A, Guilbert B;
XX Wilson IA, Knox P;
XX WPI; 2000-126380/11.
XX DR
XX Novel synthetic analogues for diagnosis of thrombosis, embolism,
XX atherosclerosis, inflammation or cancer.
XX PS Example 2; Page 17; 46pp; English.
XX CC This sequence represents an alpha-2 antipiasmin derived peptide. Alpha-2
XX antipiasmin is a substrate for the fibrin stabilising enzyme Factor
XX XIIIa. Synthetic analogues of lysine and glutamine labelled with a
XX detectable moiety can also act as substrates for Factor XIIIa. The
XX invention relates to compounds which can be used in the diagnosis of
XX thrombosis or embolism and also for diagnosing atherosclerosis,
XX inflammation or cancer. The peptide forms part of the compound
XX
XX SQ Sequence 13 AA;
XX
XX Query Match 91.0%; Score 61; DB 3; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 0.00016;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 NOEQVSPYTLTK 12
XX |||||
XX 1 NOEQVSPYTLTK 12
XX
XX Db
XX
XX RESULT 5
XX AA67841
XX ID AA67841 standard; peptide; 13 AA.
XX AC
XX AA67841;
XX XX
XX DT 25-APR-2000 (first entry)
XX XX
XX DE Alpha-2 antipiasmin enzyme derived peptide #8.
XX XX
XX KW Alpha-2 antipiasmin; fibrin stabilising enzyme; Factor XIIIa; cancer;
XX embolism; thrombosis; inflammation; atherosclerosis; diagnosis.
XX XX
XX OS Synthetic.
XX PA
XX Key Location/Qualifiers
XX Modified-site 1
XX /note= "N-terminal acetyl"
XX FT
XX Modified-site 8
XX /note= "2-Naphthylalanine"
XX FT
XX WO960018-A1.
XX PN

```

```

PD 25-NOV-1999.
XX PF 14-MAY-1999; 99WO-GB001550.
XX PR 15-MAY-1998; 98EP-00303872.
XX (AMSH ) NYCOMED AMERSHAM PLC.
XX PA
XX Storey AE, Mendizabal M, Champion S, Gibson A, Guilbert B;
XX Wilson IA, Knox P;
XX WPI; 2000-126380/11.
XX DR
XX Novel synthetic analogues for diagnosis of thrombosis, embolism,
XX atherosclerosis, inflammation or cancer.
XX PS Example 2; Page 16; 46pp; English.
XX CC This sequence represents an alpha-2 antipiasmin derived peptide. Alpha-2
XX antipiasmin is a substrate for the fibrin stabilising enzyme Factor
XX XIIIa. Synthetic analogues of lysine and glutamine labelled with a
XX detectable moiety can also act as substrates for Factor XIIIa. The
XX invention relates to compounds which can be used in the diagnosis of
XX thrombosis or embolism and also for diagnosing atherosclerosis,
XX inflammation or cancer. The peptide forms part of the compound
XX
XX SQ Sequence 13 AA;
XX
XX Query Match 88.1%; Score 59; DB 3; Length 13;
XX Best Local Similarity 92.3%; Pred. No. 0.00038;
XX Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 NOEQVSPYTLTKG 13
XX |||||
XX 1 NOEQVSPYTLTKG 13
XX
XX Db
XX
XX RESULT 6
XX AA67840
XX ID AA67840 standard; peptide; 13 AA.
XX AC
XX AA67840;
XX XX
XX DT 25-APR-2000 (first entry)
XX XX
XX DE Alpha-2 antipiasmin enzyme derived peptide #7.
XX XX
XX KW Alpha-2 antipiasmin; fibrin stabilising enzyme; Factor XIIIa; cancer;
XX embolism; thrombosis; inflammation; atherosclerosis; diagnosis.
XX XX
XX OS Synthetic.
XX PA
XX Key Location/Qualifiers
XX Modified-site 1
XX /note= "N-terminal acetyl"
XX FT
XX WO960018-A1.
XX PN
XX DT 25-NOV-1999.
XX XX
XX DE 14-MAY-1999; 99WO-GB001550.
XX PF 15-MAY-1998; 98EP-00303872.
XX PR (AMSH ) NYCOMED AMERSHAM PLC.
XX PA
XX Storey AE, Mendizabal M, Champion S, Gibson A, Guilbert B;
XX Wilson IA, Knox P;
XX WPI; 2000-126380/11.
XX DR
XX Novel synthetic analogues for diagnosis of thrombosis, embolism,
XX atherosclerosis, inflammation or cancer.
XX PT

```

```

XX Example 2; Page 16; 46pp; English.
PS
CC This sequence represents an alpha-2 antiplasmin derived peptide. Alpha-2
CC antiplasmin is a substrate for the fibrin stabilising enzyme Factor
CC XIIIa. Synthetic analogues of lysine and glutamine labelled with a
CC detectable moiety can also act as substrates for Factor XIIIa. The
CC invention relates to compounds which can be used in the diagnosis of
CC thrombosis or embolism and also for diagnosing atherosclerosis,
CC inflammation or cancer. The peptide forms part of the compound
XX
SQ Sequence 13 AA;

Query Match      88.1%; Score 59; DB 3; Length 13;
Best Local Similarity 92.3%; Pred. No. 0.00038;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NOEQVSPYTLKG 13
DB 1 NOEQVSPYTLKG 13

RESULT 7
AA67844
ID AA67844 standard; peptide; 13 AA.
XX
AC AA67844;
XX
DT 25-APR-2000 (first entry)
XX
DE Alpha-2 antiplasmin enzyme derived peptide #11.
XX
KW Alpha-2 antiplasmin; fibrin stabilising enzyme; Factor XIIIa; cancer;
KW embolism; thrombosis; inflammation; atherosclerosis; diagnosis.
XX
OS Synthetic.
XX
PN W09960018-A1.
XX
PD 25-NOV-1999.
XX
PF 14-MAY-1999; 99WO-GB001550.
XX
PR 15-MAY-1998; 98EP-00303872.
XX
PA (AMSH ) NYCOMED AMERSHAM PLC.
XX
PI Storey AE, Mendizabal M, Champion S, Gibson A, Guilbert B;
PI Wilson IA, Knox P;
XX
DR WPI; 2000-126380/11.
XX
PT Novel synthetic analogues for diagnosis of thrombosis, embolism,
PT atherosclerosis, inflammation or cancer.
XX
PS Example 2; Page 17; 46pp; English.
CC This sequence represents an alpha-2 antiplasmin derived peptide. Alpha-2
CC antiplasmin is a substrate for the fibrin stabilising enzyme Factor
CC XIIIa. Synthetic analogues of lysine and glutamine labelled with a
CC detectable moiety can also act as substrates for Factor XIIIa. The
CC invention relates to compounds which can be used in the diagnosis of
CC thrombosis or embolism and also for diagnosing atherosclerosis,
CC inflammation or cancer. The peptide forms part of the compound
XX
SQ Sequence 13 AA;

Query Match      68.7%; Score 46; DB 3; Length 13;
Best Local Similarity 69.2%; Pred. No. 0.11;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 NOEQVSPYTLKG 13
DB 1 NOEQVSPYTLKG 13

```

```

DB 1 NOEQVSPYAAAAG 13

RESULT 8
AAM50298
ID AAM50298 standard; peptide; 13 AA.
XX
AC AAM50298;
XX
DT 04-FEB-2002 (first entry)
XX
DE Factor XIIIa substrate-polylysine peptide chimera.
XX
KW Tissue engineering; tissue repair; tissue regeneration; drug delivery;
KW Factor XIIIa; heparin.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FH Modified-site 1
FH /note="dansyl leucine, allows detection of peptide by
FT fluorescence"
FT Peptide 2..8
FT /note="Factor XIIIa substrate"
FT Peptide 9..13
FT /note="heparin-binding peptide"
XX
PN W0200183522-A2.
XX
PD 08-NOV-2001.
XX
PF 01-MAY-2000; 2000WO-US011947.
XX
PR 01-MAY-2000; 2000WO-US011947.
XX
PA (BIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.
PA (HUBB/) HUBBELL J A.
XX
PI Hubbell JA, Schense JC, Sakiyama-Elbert SE;
XX
DR WPI; 2002-041479/05.
XX
PT Novel growth factor modified protein matrix useful for tissue repair,
PT regeneration and/or remodeling, and/or drug delivery.
XX
PS Example 2; Page 25; 51pp; English.
XX
CC The present sequence is that of a fluorescent synthetic peptide, which
CC has a Factor XIIIa substrate at the N-terminal end and a polylysine
CC moiety at the C-terminus. The peptide is used to produce a heparin-
CC peptide chimera by coupling to a heparin oligosaccharide, with a unique
CC aldehyde group on one end, via reductive amination. A matrix is formed in
CC which heparin is used to incorporate heparin-binding growth factors for
CC use in tissue repair, regeneration, remodeling, and/or drug delivery. The
CC invention provides biocompatible and biodegradable matrices incorporating
CC bioactive proteins such as growth factors. The matrices can be formed in
CC vitro or in vivo at the site of implantation, and provide controlled
XX
SQ Sequence 13 AA;

Query Match      53.7%; Score 36; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NOEQVSP 7
DB 2 NOEQVSP 8

RESULT 9
ABG73917
ID ABG73917 standard; peptide; 13 AA.

```



```

XX AC ABG73917;
XX XX
XX DT 27-MAR-2003 (first entry)
XX DE Factor IIIa cleavage site-containing peptide #1.
XX KW Factor XIIIa; fibrin matrix; heparin; heparin binding protein; HBP;
XX KM wound healing; vasculature; skin; nerve; liver.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1
XX FT /label= OTHER
XX FT /note= "Dansyl-leucine, used to allow detection by
XX FT fluorescence"
XX PN US6468731-B1.
XX XX
XX PD 22-OCT-2002.
XX XX
XX PF 29-SEP-2000; 2000US-00675922.
XX PR 27-AUG-1998; 98US-00141153.
XX XX
XX PA (EIDG-) EIDGENOESSISCHE TECH HOCHSCHULE ZUERICH.
XX PA (UYZU-) UNIV ZUERICH.
XX PI Hubbell JA, Schense JC, Sakiyama SE;
XX XX
XX DR WPI; 2003-155552/15.
XX XX
XX PT Novel matrix for controlled delivery of heparin binding protein has
XX PT fibrin, peptide having first domain with Factor XIIIa substrate and
XX PT second domain with heparin/heparin-like compound, and heparin binding
XX PT protein.
XX PS Example 2; Col 6; 17pp; English.
XX XX
XX CC The invention relates to a matrix comprising fibrin, a peptide which
XX CC comprises a first domain comprising a Factor XIIIa substrate and a
XX CC second domain comprising heparin or a heparin-like compound coupled to
XX CC the first domain, and a heparin binding protein (HBP), where the peptide
XX CC is covalently attached to the fibrin through the first domain. The matrix
XX CC is used for controlled release of HBP, for enhanced wound healing in
XX CC various tissues including vasculature, skin, nerve and liver. Heparin
XX CC used in the matrix protects the growth factors from proteolytic
XX CC degradation and prolong their activity until they are released from the
XX CC matrix. Despite their relatively strong affinity for heparin, heparin
XX CC binding growth factors dissociate from the matrix on a short time scale.
XX CC Therefore, high excess of binding sites is essential to ensure that they
XX CC do not diffuse far before they bind to the matrix again. This equilibrium
XX CC also allows for the binding of free growth factor to cell surface
XX CC receptors that are in close proximity to the site of dissociation. This
XX CC method of controlled release provides both relatively long term binding
XX CC of growth factors and rapid release of growth factors to local cells. The
XX CC present sequence is a peptide used to make the matrix of the invention
XX XX
XX SQ Sequence 13 AA;
XX
XX Query Match 53.7%; Score 36; DB 6; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 8.2;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

XX AC AAY20946;
XX XX
XX DT 22-JUL-1999 (first entry)
XX DE Human prenasilin II mutant protein fragment 39.
XX KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
XX KW frameshift mutation; age-related disease; neurodegenerative disorder;
XX KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
XX KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
XX KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
XX KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
XX KW neurofilament-F; prenasilin I; prenasilin II; cellular tumour antigen;
XX KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
XX KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
XX KW high mobility group protein-C; neuroendocrine specific protein A.
XX XX
XX OS Synthetic.
XX OS Homo sapiens.
XX PN WO9845322-A2.
XX XX
XX PD 15-OCT-1998.
XX XX
XX PF 02-APR-1998; 98WO-IB000705.
XX PR 10-APR-1997; 97US-0043163P.
XX XX
XX PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
XX PA (UYRO-) UNIV ROTTERDAM ERASMUS.
XX PA (UYUT-) RIJKSUNIV UTRECHT.
XX XX
XX PI Van Leeuwen FW, Grosveld FG, Burbach JPH;
XX XX
XX DR WPI; 1998-609901/51.
XX DR N-PSDB; AAY75762.
XX XX
XX PT Diagnosing disease by detecting frameshift mutations in RNA or
XX PT corresponding protein mutations - used to diagnose cancer and
XX PT neurological diseases, particularly Alzheimer's disease, and also for
XX PT treatment and prevention with specific ribozymes or wild-type RNA.
XX PS Disclosure; Fig 11; 258pp; English.
XX XX
XX CC This invention describes a novel method for the diagnosis of a disease
XX CC caused by, or associated with, an RNA molecule that has a frameshift
XX CC mutation. The method is used to diagnose age-related diseases, especially
XX CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
XX CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II
XX CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
XX CC and many others listed) or susceptibility to these disorders. The method
XX CC allows a definitive diagnosis of Alzheimer's disease in living patients,
XX CC at an early stage. It is based on the observation that disease may be
XX CC caused by mutations in RNA rather than DNA. The invention describes the
XX CC use of neuronal system RNA molecules, specifically proteins including
XX CC beta-amyloid precursor protein (beta-APP), the microtubule associated
XX CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
XX CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
XX CC neurofilament-F, prenasilin I, prenasilin II, glial fibrillary acidic
XX CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
XX CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
XX CC protein-C (HMGP-C) and neuroendocrine specific protein A
XX XX
XX SQ Sequence 13 AA;
XX
XX Query Match 40.3%; Score 27; DB 2; Length 13;
XX Best Local Similarity 44.4%; Pred. No. 4.1e+02;
XX Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

```

RESULT 10
AAY20946
ID AAY20946 standard; protein; 13 AA.

```



```

FH Key Location/Qualifiers
FT Region 6..9
XX /note= "PPPPY motif"
XX
XX WO9737223-A1.
XX
XX 09-OCT-1997.
XX
XX 03-APR-1997; 97WO-US005547.
XX
XX 03-APR-1996; 96US-00630916.
XX
XX (CYTO-) CYTOGEN CORP.
XX (UNNC-) UNIV NORTH CAROLINA.
XX
XX Pirozzi G, Kay BK, Fowlkes DM;
XX
XX WPI, 1997-503234/46.
XX
XX Identifying cell signalling and growth regulatory polypeptides by
XX reaction with multivalent recognition complex - polypeptides are useful
XX in targeted drug selection.
XX
XX Claim 92; Fig 27; 220pp; English.
XX
XX Peptides AAW38005-08 are the sequences of WW domain binding clones
XX obtained by screening random or biased libraries with the WW domains of
XX the novel WWP proteins. The present peptide was obtained after screening
XX with WW domain 1 of WWP1 (WWP1.1). The peptides are peptide recognition
XX units ("ligands"), and are used for analysing specificities of the WW
XX domains. The WW domain is a small functional domain. Its name is derived
XX from the observation that two tryptophan residues, one in the amino
XX terminal portion of the WW domain and one in the carboxyl terminal
XX portion, are conserved. Most proteins containing WW domains have a
XX function involving cell signalling and growth regulation or the
XX organisation of the cytoskeleton. Polypeptides containing a WW domain are
XX identified by treating a multivalent recognition unit complex that has
XX selective binding affinity for a WW domain, with many polypeptides and
XX identifying those with selective affinity for the complex. Proteins
XX containing WW domains are used for targeted drug screening, i.e. to
XX identify potential modulators of specific WW domain interactions. note:
XX sequence in SEQ ID listing differs from sequence given in figure. The SEQ
XX ID sequence is as follows: GPSRQPPPYRYTVK
XX
XX Sequence 13 AA;
XX
XX Query Match 37.3%; Score 25; DB 2; Length 13;
XX Best Local Similarity 50.0%; Pred. No. 9.8e+02;
XX Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX
XX 3 EQVSPYTLK 12
XX |||||
XX 4 EQPPYRYTVK 13
XX
XX RESULT 14
XX AAE23145
XX ID AAE23145 standard; peptide; 13 AA.
XX
XX AAE23145;
XX
XX 21-AUG-2002 (first entry)
XX
XX Influenza A virus subtype 16 haemagglutinin (HA) protein fragment #9.
XX
XX Haemagglutinin; HA protein; vaccine; infection.
XX
XX Influenza A virus.
XX
XX WC200224734-A2.
XX
XX 28-MAR-2002.
XX

```

```

PF 19-SEP-2001; 2001WO-IB001914.
XX
XX 19-SEP-2000; 2000GB-00022969.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Fouchier R, Olsen B, Osterhaus A;
XX
XX WPI, 2002-416272/44.
XX
XX
XX New isolated protein of influenza A virus subtype H16 useful for
XX preparing vaccines for treating influenza A virus infection, and for
XX detecting influenza virus in a sample.
XX
XX Claim 4; Page 57; 59pp; English.
XX
XX The invention relates to influenza A virus subtype (designated H16)
XX haemagglutinin (HA) proteins and nucleic acid encoding such proteins.
XX Sequences of the invention are useful as vaccines. They are also used in
XX the manufacture of a medicament for treating or preventing influenza A
XX virus infections. The present sequence is influenza A virus subtype H16
XX HA protein fragment
XX
XX Sequence 13 AA;
XX
XX Query Match 37.3%; Score 25; DB 5; Length 13;
XX Best Local Similarity 66.7%; Pred. No. 9.8e+02;
XX Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
XX 6 SPYTL 11
XX :|||||
XX 2 NPYTLV 7
XX
XX Db
XX
XX RESULT 15
XX AAR34779
XX ID AAR34779 standard; peptide; 13 AA.
XX
XX AAR34779;
XX
XX 25-MAR-2003 (revised)
XX 16-JUL-1993 (first entry)
XX
XX CAD tryptic peptide, internal sequence 10.
XX
XX Tobacco; cinamy] alcohol dehydrogenase; CAD; tryptic peptide; subunit;
XX probe; primer; plant; lignin; biosynthesis; biosynthesis; digestibility;
XX feed; cellulose; crop; pathogen; timber.
XX
XX Nicotiana tabacum.
XX
XX WO9305159-A1.
XX
XX 18-MAR-1993.
XX
XX 27-APR-1992; 92WO-GB000774.
XX
XX 26-APR-1991; 91GB-00009063.
XX
XX (ICIL ) IMPERIAL CHEM IND PLC.
XX
XX Boudet AM, Inze DG, Schuch WW;
XX
XX WPI, 1993-100982/12.
XX
XX Modifying lignin content of plants - comprises use of recombinant DNA
XX encoding lignin biosynthesis enzyme, e.g. CAD.
XX
XX Disclosure; Fig 1; 69pp; English.
XX
XX The sequences given in AAR34768-82 represent peptide sequences derived
XX from tobacco cinamy] alcohol dehydrogenase (CAD) tryptic peptides.
XX Purified CAD consists of two subunits of approximate molecular weights
XX

```

CC 42.5 kD and 44 kD. These peptides were derived by trypsin digestion of  
CC both polypeptides. These peptides were used to produce probes and primers  
CC which were used in the isolation of the CAD cDNA. The CAD cDNA may be  
CC used in the production of a recombinant DNA comprising a gene which is  
CC essential to plant lignin biosynthesis. When incorporated into a plant  
CC genome by transformation, mRNA transcribed from the coding region  
CC inhibits production of the enzyme from the endogenous gene. The  
CC recombinant DNA may be used to provide plants having altered ability to  
CC synthesise lignin. The principle applications of altered lignin  
CC biosynthesis are improvement of digestibility of animal forage crops,  
CC reduction of lignin in woody feedstocks for cellulose fibre extraction,  
CC improvement of the response of crop plants to pathogen attack and  
CC improvement of timber quality. (Updated on 25-MAR-2003 to correct PN  
CC field.)  
XX

SQ Sequence 13 AA;

Query Match 35.8%; Score 24; DB 2; Length 13;  
Best Local Similarity 80.0%; Pred. No. 1.5e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 VSPLYT 9  
: |||  
Db 5 TSPYT 9

Search completed: October 1, 2004, 13:47:59  
Job time : 54 secs